

---

# **SYNTHESIS AND STRUCTURAL CHARACTERIZATION OF NEW OLIGONUCLEOTIDE AND NUCLEOSIDE ANALOGUES AS PHARMACOLOGICAL TOOLS**

---

**Jussara Amato**

Dottorato in Scienze Biotechnologiche – XIX ciclo  
Indirizzo Biotechnologie Molecolari  
Università di Napoli Federico II





Dottorato in Scienze Biotecnologiche – XIX ciclo  
Indirizzo Biotecnologie Molecolari  
Università di Napoli Federico II



---

# **SYNTHESIS AND STRUCTURAL CHARACTERIZATION OF NEW OLIGONUCLEOTIDE AND NUCLEOSIDE ANALOGUES AS PHARMACOLOGICAL TOOLS**

---

**Jussara Amato**

Dottoranda: Jussara Amato

Relatore: Prof. Gennaro Piccialli

Coordinatore: Prof. Gennaro Marino



## ***DEDICATIONS***

Dedicated to the most important and inspirational people in my life:

**My dad**

from whom I inherited the “chemistry gene”.



# INDEX

## ABSTRACTS

Short abstract	pag. i
Long abstract	pag. ii

## CHAPTER 1

### *“General Introduction”*

1.1 Synthesis and structural characterization of oligonucleotide (ON) analogues	pag.1
1.2 Synthesis of nucleoside and nucleotide analogues	pag.3

## CHAPTER 2

### *“Synthesis of TFO (Triplex Forming Oligonucleotides) containing a 3'-3' polarity inversion site conjugated to moieties (acridine residue) capable to stabilize the triple helix complexes or to supply particular biological activities”*

Introduction	pag. 6
2.1. DNA triple helix formation	pag.8
2.1.1. Hoogsteen base-pairing	pag.9
2.1.2. Natural Hoogsteen-bonded triplexes	pag.9
2.2 pH and cation effects on the properties of DNA triplexes	pag.10
2.2.1. Protonation effect on the triplex thermodynamic	pag.10
2.2.2. Ionic strength effects on triplex stability	pag.10
2.2.3 Cation effect on the triplex thermodynamic	pag.11
2.3. Triplex applications	pag.11
2.3.1. Transcription	pag.12
2.3.2. Genome modification	pag.12
2.3.3. Triplex-directed mutagenesis	pag.12
2.3.4. Triplex-induced recombination	pag.13
2.3.5. Other applications: properties of functional molecules conjugates-TFOs	pag.14
2.4. Recent improvements in antigene technology	pag.15
2.4.1. Increasing triplex stability	pag.15
2.4.2. Dual recognition	pag.15
2.4.3. Alternate-Strand Triple helix formation	pag.17
2.5. The aim of the work	pag.19
2.6. Chemistry	pag.20
2.7. UV and circular dichroism (CD) studi	pag.23
2.8. Conclusions	pag.25

<b>2.9. Experimental Session</b>	pag.25
2.9.1. <i>General Methods</i>	pag.25
2.9.2 <i>General Procedures</i>	pag.25
<b>References</b>	pag.27

## CHAPTER 3

### ***“Synthesis and characterization of new kind ON analogues able to form stable monomolecular quadruple helices DNA complexes: Tetra-End-Linked Oligonucleotides (TEL-ON)”***

<b>Introduction</b>	pag.29
<b>3.1. Structural polymorphism of G-quadruplex DNA</b>	pag.30
3.1.1 <i>Topological classification of G-Quadruplex DNA structures</i>	pag.30
3.1.2. <i>G-Quadruplex DNA groove widths</i>	pag.32
3.1.3. <i>G-Quadruplex DNA metal ion binding</i>	pag.33
<b>3.2. Other tetrads</b>	pag.35
<b>3.3. Biological relevance of G-quadruplex DNA</b>	pag.35
<b>3.4. G-quadruplex ligands: molecular recognition</b>	pag.36
3.4.1. <i>Molecular Recognition of G-Quadruplexes by Proteins</i>	pag.37
3.4.2. <i>Quadruplex Aptamers : Recognition by Small Molecules</i>	pag.38
3.4.3. <i>Small Molecules that Stabilize DNA G-Quadruplexes: Telomerase Inhibitors</i>	pag.38
<b>3.5. Guanosine self-assembly in materials science, biosensor design, and nanotechnology</b>	pag.40
3.5.1. <i>DNA Nanostructures: G-Wires, Frayed Wires, and Synapses</i>	pag.40
3.5.2. <i>Toward Synthetic Ion Channels</i>	pag.40
3.5.3. <i>Formation of Biosensors and Nanomachines with G-Quadruplex DNA</i>	pag.40
<b>3.6. Future direction</b>	pag.41
<b>3.7. The aim of the work</b>	pag.41
<b>3.8. Synthesis and characterization of a new kind ON analogues able to form stable monomolecular quadruple helices DNA complexes (TEL-ON)</b>	pag.43
<b>3.9. Studies on the length of tetra-end-linked linker on the orientation of the ON strands aimed at of obtaining stable quadruplex structures</b>	pag.55
<b>3.10. Studies on quadruplex complexes, containing less stable quartets such as T-tetrads</b>	pag.60
<b>3.11. Conclusions</b>	pag.66
<b>References</b>	pag.67



## CHAPTER 4

### ***“Synthesis of combinatorial libraries of nucleosides, specifically inosine analogues”***

Introduction	pag.69
4.1. Nucleoside analogues	pag.71
4.2. Nucleoside analogue uptake	pag.72
4.3 Nucleoside analogue activation	pag.73
4.4. Mechanism of action of nucleoside analogues	pag.73
4.5. Mitochondrial toxicity of nucleoside analogues	pag.75
4.6. Nucleotide analogues	pag.75
4.7. Solid-phase synthesis of nucleoside analogues	pag.76
4.8. The aim of the work	pag.77
4.9. Chemistry	pag.77
4.10. Conclusions	pag.80
4.11 Experimental session	pag.80
4.11.1. General Methods	pag.80
4.11.2 General Procedures	pag.81
References	pag.85

## CHAPTER 5

### ***“Synthesis of a new series of cIDPR (cyclic hypoxanthine-diphospho-ribose) analogues, designed as novel stable mimics of cADPR a potent $Ca^{2+}$ -mobilizing second messenger”***

Introduction	pag.86
5.1. Cyclic-ADP-Ribose: a new way to control calcium	pag.88
5.2. ADP-ribosyl cyclase	pag.89
5.3. CD 38	pag.90
5.4. Stability of cADPR	pag.91
5.5. Synthesis of cADPR analogues	pag.92
5.5.1. Enzymatic synthesis of cADPR derivatives	pag.93
5.5.2. Non-enzymatic syntheses of cADPR analogues	pag.95
5.6. Chemical cyclization strategies for new cADPR analogues synthesis	pag.98
5.7. Structural cADPR analogues for SAR elucidation	pag.100
5.8 The aim of the work	pag.101

<b>5.9. Chemistry</b>	pag.102
<b>5.10 Conclusions</b>	pag.104
<b>5.11 Experimental Session</b>	pag.104
5.11.1. General Methods	pag.104
5.11.2. General Procedures	pag.106
<b>References</b>	pag.111
<b>Appendix</b>	pag.112
<i>List of Publications and Communications</i>	pag.112

## **ACKNOWLEDGEMENTS**

I would like to acknowledge all the wonderful people who provided support through my entire PhD experience. Especially my mentor Professor Gennaro Piccialli, who has always been a pillar of support for all my endeavours and for putting up with my endless questions and requests, no matter how crazy. To all members of the laboratory, including Dr.Nicola Borbone and Dr.Stefano D'Errico, particularly to Dr.Giorgia Oliviero I would like to say, "I don't know what I would do without you!" My dad, Mr.Francesco Amato, mom, Mrs.Antonina Guercio, who motivate me to keep going with their love. My brothers Arturo for his help with many of my thesis figures and Vincenzo for constantly asking me when I am going to take one's PhD. My boyfriend, Alessandro Leo, who definitely has the patience of an angel; I know because I keep testing him over and over. My Naples co-residents Miss Marica De Gruttola and Fabiana Molinario for their companionships and especially thank to Miss Loredana Molinario for providing the moral support to actually achieve that goal. My friends, including Miss Filomena Pierri, who was a ready supply of humour, caffeine and "nutella" when I needed it the most. And last but not least, the professors Luciano Mayol and Lorenzo De Napoli, the department of Natural Compounds Chemistry (especially Dr. Luisa Cuorvo) and the Biotechnology-PhD, that provide me the opportunity of fees due in this field. In these past months and years I have had both wonderful and bad experiences in my life, and ALL of you made those experiences very memorable by always being there for me.

Sincerely,  
Jussara Amato

## Short Abstract

Nucleoside, nucleotide and nucleic acid researches have been attracting much attention in recent years for their useful and promising biological and therapeutical application. This class of molecules results in efficient tools not only as “molecular sensors”, but also to modulate well identified cellular activities, introducing an interdisciplinary approach going from the biomedical and biotechnological research to the clinical treatment of many diseases, until technological applications in nanoscience field.

Synthetic oligonucleotides (ON) of 15-17 bases having characteristic sequences can be efficient gene expression inhibitors or, more generally, regulators, selectively interfering with the normal biological functions of nucleic acids or of specific proteins. Oligonucleotides, addressed to the control of a specific gene expression, can be catalogued, on the base of their target and mechanism of action, in “antisense” or “antigene”, when they target a specific mRNA or a tract of duplex DNA, respectively, or as “aptamers”, when selectively recognize a specific protein.

The first part of my research work has been addressed to synthesize and study two main classes of ON analogues, with potential biological properties:

- 1) ON analogues able to form stable triple helices
- 2) ON analogues able to form new monomolecular quadruplex complexes

Then, special attention has been devoted to structural investigation on the complexes formed by the synthesized oligonucleotides.

Pharmacological structural investigations on complexes formed by modified ONs and synthetic models of their natural targets (mRNA, DNA or proteins), essentially carried out by NMR spectroscopy associated with mechanic and molecular dynamics calculations, allow a better comprehension of specific recognition mechanisms. This is useful to design new rational structural modifications to be introduced in synthetic ONs. Recently, medicinal chemistry research emphasizes the need of nucleoside and nucleotide analogues in high number and diversity, to be used as building blocks in the synthesis of new ON generations. Apart from being the genomic building blocks, nucleosides interact with roughly one-third of protein classes in the human genome, all targets of therapeutic importance in biological systems. Many nucleoside and nucleotide synthetic analogues have been the cornerstone of antiviral therapy over the past 30 years, and many others also exhibit anti-proliferative, antibiotic and antifungal properties. The rapid development of drug resistance and high toxicity that characterize the use of these agents, suggest the novel nucleoside drug discovery for critical medical needs. The availability of high-throughput screening capabilities together with the combinatorial synthesis of small organic molecule libraries offers a unique opportunity to accelerate the discovery in these fields.

The usefulness of mimics in this field is not only as a pharmacological tool, but also in the elucidation of the mechanism of natural analogues action, labelling of binding proteins or antagonizing of their activities.

In these contexts, the second part of my research has been employed and addressed on such targets:

- 3) Synthesis of combinatorial libraries of nucleosides, specifically inosine analogues
- 4) Synthesis of new cIDPR (cyclic hypoxanthine-diphospho-ribose) analogues, designed as novel stable mimics of cADPR, a potent  $\text{Ca}^{2+}$ -mobilizing second messenger.

## Long Abstract

Il principale obiettivo delle ricerche in campo farmaceutico-farmacologico è l'ottenimento di un farmaco che sia in grado di interagire col suo bersaglio in modo specifico o, almeno, altamente selettivo. La progettazione di un tale farmaco su base razionale richiede una conoscenza dettagliata del suo bersaglio, che non sempre è disponibile nel caso in cui esso sia una proteina. La situazione è certamente più favorevole quando il bersaglio è un DNA o un RNA perché la struttura e le funzioni di tali classi di molecole sono meglio conosciute nelle grandi linee e, soprattutto, perché sono certamente più chiari i meccanismi di riconoscimento molecolare in cui esse sono coinvolte. Perché una molecola, quale potenziale farmaco, sia in grado di interagire specificamente con un particolare acido nucleico, deve essere capace di "leggere" una sequenza non più breve di 16-18 basi. L'unica molecola capace di fare ciò è un altro acido nucleico. Per questo motivo, nell'ultimo decennio l'interesse verso gli oligonucleotidi sintetici (ON) è cresciuto. Gli ON deputati a queste funzioni, possono essere catalogati in base al tipo di bersaglio e al loro meccanismo d'azione. Essi possono agire come molecole "antisense" o "antigene", i cui target sono, rispettivamente, uno specifico mRNA ed un tratto di DNA a doppia elica, o come aptameri, cioè riconoscere in maniera assolutamente selettiva una proteina.

La strategia antisense si basa sul riconoscimento specifico della sequenza bersaglio di uno specifico mRNA da parte di un ON avente sequenza complementare mediante formazione di un ibrido DNA-RNA stabilizzato da legami Watson-Crick<sup>[1]</sup>.

La formazione di tale ibrido porta, attraverso vari meccanismi, al blocco della sintesi della relativa proteina. Nel secondo caso, l'ON agisce ancora più a monte, avendo come bersaglio un tratto di DNA con il quale forma una tripla elica legandosi nel solco maggiore del DNA duplex. La formazione di questi complessi localizzati a tripla elica può impedire le interazioni dei geni con i necessari fattori di trascrizione o fisicamente bloccare l'iniziazione o l'elongazione del complesso di trascrizione<sup>[2]</sup>.

Il processo di riconoscimento molecolare alla base di questa strategia è estremamente selettivo, essendo assicurato dalla formazione di triplette di basi stabilizzate da legami idrogeno, secondo schemi di tipo Hoogsteen o reverse Hoogsteen<sup>[3]</sup>. Oltre alla elevata specificità d'azione, un ON rappresenta un farmaco ideale dal punto di vista farmacodinamico, perché può agire ad una concentrazione cellulare molto bassa. Tuttavia, l'impiego terapeutico è limitato da diversi problemi farmacocinetici, come la loro scarsa capacità di penetrare le membrane cellulari e la veloce degradazione ad opera delle nucleasi cellulari. Per ovviare a questi inconvenienti, sono stati progettati e sintetizzati vari ON analoghi e, dopo circa un ventennio di studi, si è oggi arrivati all'approvazione dei primi ON antisense come farmaci antivirali, mentre altri sono in avanzata fase di sperimentazione clinica come antitumorali. Le applicazioni antigene sono ulteriormente limitate dal fatto che, sul DNA duplex bersaglio, è necessaria la presenza di tratti omopurinici-omopirimidinici lunghi almeno 16-18 coppie di basi. Per ampliare il numero di sequenze di DNA riconoscibili mediante formazione di strutture triplex, Horne e Dervan<sup>[4]</sup>, seguiti poi da altri gruppi di ricerca, hanno proposto l'utilizzazione di ON contenenti un'inversione di polarità, per il riconoscimento di duplex a sequenza mista del tipo (purina)<sub>m</sub>(pirimidina)<sub>n</sub>.

Un'ulteriore applicazione terapeutica degli ON sintetici è costituita dal loro potenziale impiego come aptameri<sup>[5]</sup>. Un aptamero è un ON in grado di assumere una conformazione che gli consenta di interagire specificamente con una particolare proteina, modulandone la funzione. La dimostrazione che corti frammenti di DNA o di RNA sono in grado di legare specificamente alcune proteine, fa sì che essi non vengano più considerati solo come

trasportatori dell'informazione genetica ma anche come biomolecole coinvolte in diversi processi cellulari. Gli ON aptameri, sono acidi nucleici a doppio o a singolo filamento capaci di legare e quindi inibire particolari proteine coinvolte nei processi di controllo genico. In questa classe rientrano anche gli ON con sequenze poli-G, formanti quadruple eliche di DNA caratterizzate dalla strutturazione di quartetti di guanine. Il crescente interesse verso tali strutture è legato alla scoperta dell'attività inibitoria che alcuni ON sintetici, in grado di formare quadruplex, mostrano nei confronti di alcune importanti proteine quali la trombina e l'integrasi dell'HIV-1<sup>[6-7]</sup>, quest'ultima responsabile dell'integrazione del DNA virale in quello della cellula ospite. L'importanza e l'attualità degli studi in questo campo deriva anche dal fatto che, sequenze telomeriche del gene, ricche di deossiguanosine, possono formare complessi a quadrupla elica<sup>[8]</sup>. Queste strutture, probabilmente presenti all'estremità 3' dei telomeri degli eucarioti e dei procarioti, potrebbero giocare un ruolo importante nella regolazione dell'attività telomerasica. L'espressione ubiquitaria della telomerasi nelle cellule tumorali e la sua ridotta attività nelle cellule normali ha suggerito l'ipotesi del coinvolgimento di questo enzima nel processo di carcinogenesi. E' stato ipotizzato che le sequenze telomeriche del gene, strutturandosi in quadrupla elica (quadruplex), riescono a regolare l'attività della telomerasi<sup>[9]</sup>. In tale ambito le G-quadruplex risultano utili aptameri<sup>[5]</sup> per le proteine che regolano la lunghezza dei telomeri dei geni nelle cellule eucariote, influenzando i meccanismi di senescenza e carcinogenesi cellulare. Di qui, il duplice interesse farmacologico verso le strutture quadruplex del DNA, inibizione specifica di alcune proteine (funzionamento come aptamero) o bersaglio per l'azione di farmaci antitumorali<sup>[10]</sup>; inoltre esse rappresentano validi modelli di studio per le interazioni aptamero-proteina. Recenti studi ne hanno inoltre evidenziato un'applicazione "non farmacologica" suggerendo che strutture G-quadruplex possano essere utilizzate come micro-conduttori di elettricità, grazie alla capacità di trasferire elettroni lungo la quadrupla elica<sup>[11-12]</sup>. Inoltre, queste strutture, in ragione della capacità di selezionare particolari conformazioni e della loro elevata stabilità, sembrano ottime candidate nella costruzione di nano-materiali e nano-meccanismi, questi ultimi basati su cambi conformazionali controllati e ripetitivi. Tutte queste promettenti possibilità di utilizzo degli ON rendono necessaria la progettazione e la sintesi di nuove generazioni di ON opportunamente modificati, allo scopo di individuare nuove e più efficaci applicazioni in campo biochimico e biomedico. Studi strutturali, eseguiti mediante tecniche di risonanza magnetica nucleare (NMR) associate a calcoli di meccanica e dinamica molecolare, sugli ON naturali e/o modificati e sui vari complessi che questi possono formare con i relativi bersagli, permettono una migliore comprensione dei meccanismi di riconoscimento che regolano le interazioni tra queste molecole. Le informazioni, così ottenute, possono essere utilizzate per una più razionale progettazione delle modifiche strutturali da introdurre sugli ON sintetici, al fine di ottimizzarne le proprietà biologiche. In tale ambito, la ricerca medica più recente enfatizza il bisogno di sintetizzare nucleosidi e nucleotidi in elevato numero e diversità, per poter realizzare una vasta collezione di ON opportunamente modificati. Più che la sintesi classica, sembra che le metodiche sintetiche combinatoriali in fase solida possano soddisfare tale esigenza. Questa strategia permette, inoltre, di ottenere classi di molecole che differiscono in maniera puntiforme intorno ad una struttura centrale con potenziale attività farmacologica, rivestendo notevole interesse nella realizzazione di nuovi farmaci antivirali e antitumorali. L'elevata rapidità di sviluppo di fenomeni di resistenza dei virus, nonché la tossicità che si registra in concomitanza all'impiego di antivirali e antitumorali, sono indicativi di un crescente bisogno di ottenere nuove molecole in tale ambito. L'enorme versatilità dei nucleosidi, legata non solo al fatto che costituiscono le unità base del patrimonio genetico, ma anche alla loro capacità di interagire con più di un terzo delle proteine umane, tra cui polimerasi, chinasi, reduttasi,

recettori di membrana e proteine strutturali, spiega il notevole bisogno di avere nuovi analoghi nucleosidici, terapeuticamente attivi contro tutti questi target biologici. La natura stessa offre diversi esempi di analoghi nucleosidici capaci di agire selettivamente come inibitori della sintesi proteica (puromicina), inibitori della glicosil transferasi (tunicamicina), ed inibitori della metil transferasi (sinefungina).

Da più di 30 anni, analoghi sintetici dei nucleosidi trovano impiego nella terapia delle infezioni virali. Attualmente, i nucleosidi e nucleotidi analoghi presenti sul mercato sono circa 50, mentre più di un'ottantina sono in fase di studio (pre) clinico non solo come antivirali ma anche con attività antitumorale, antifungina, etc. Tra questi annoveriamo l'arabinosilcitidina e la 5-aza-citidina utilizzate per il trattamento di leucemie e la 5-iodouridina utilizzata per le infezioni da virus Herpes Simplex; inoltre l'aciclovir, un farmaco antivirale appartenente alla classe degli aciclonucleosidi; la 3'-tiocitidina (3TC), inibitrice della trascrittasi inversa virale, usata contro le infezioni di epatite B; ed il 3'-azido-3'-deossitimidina (AZT) il più importante nucleoside anti-HIV. Sebbene la ricerca sui nucleosidi modificati sia fortemente finalizzata alla possibilità di una loro applicazione diretta nella pratica clinica, non è da trascurare il fatto che numerosi nucleosidi e nucleotidi svolgono anche un ruolo centrale nel metabolismo cellulare, come modulatori endogeni delle funzioni cellulari. Ciò presuppone la possibilità di utilizzo di tali molecole come utili modelli per meglio comprendere i meccanismi relativi a processi biochimici a cui prendono parte analoghi nucleosidici naturali. Un esempio è rappresentato dall'adenosina - 5' - difosfato - ribosio - ciclico (cADPR), un metabolita endogeno intimamente coinvolto nella regolazione della concentrazione dello ione calcio intracellulare  $[Ca^{+2}]_i$ , in misura maggiore dell' inositolo-1,4,5-trifosfato ( $IP_3$ ) ed in maniera indipendente da esso. Poiché lo ione  $Ca^{2+}$  è coinvolto nella regolazione di diverse funzioni cellulari, come la contrazione muscolare, la secrezione di neurotrasmettitori, di ormoni ed enzimi responsabili della fertilizzazione degli ovociti, nonché nell'attivazione e proliferazione dei linfociti  $T^{[13]}$ , la cascata dei segnali modulata dal cADPR diventa un interessante target farmacologico. L'interesse verso il cADPR non è limitato solo alle sue potenzialità farmacologiche, esso infatti sembra essere utile nella delucidazione del suo meccanismo d'azione, nello screening delle proteine coinvolte nel suo metabolismo e nella sua attività regolatoria, nonché nello studio e nel *design* di analoghi del cADPR. Tutto questo potenziale *pool* di applicazioni risulta inficiato dalla sua scarsa stabilità in condizioni fisiologiche ( $t_{1/2}$  = 24h). Il cADPR è caratterizzato da un legame N-1-glicosidico fortemente labile, suscettibile a rapida idrolisi sia enzimatica, ad opera della cADPRibosio-idrolasi, sia non-enzimatica, con formazione del prodotto di idrolisi ADP-ribosio, già in soluzione acquosa neutra<sup>[14]</sup>. Tuttavia il suo importante coinvolgimento biologico rappresenta uno stimolo forte verso la progettazione e la sintesi di nuovi analoghi sufficientemente stabili e farmacologicamente utili. In questo contesto si inserisce il mio programma di ricerca svolto nei tre anni di dottorato.

In sintesi, la mia attività di ricerca è stata indirizzata ed ha prodotto risultati nel campo della :

- **Sintesi di librerie combinatoriali di nucleosidi (nella fattispecie analoghi dell'inosina).**

Obiettivo di questo lavoro di ricerca è stato la realizzazione di una metodologia sintetica versatile ed efficace mirata, in un primo momento, all'ottenimento di una collezione di analoghi N-1-alchilati dell' inosina e poi estesa ai corrispondenti 2',3'-seco-N1-alchil-inosina derivati, i quali, in virtù dell'analogia strutturale con composti a riconosciuta attività antivirale, quali il ganciclovir, il penciclovir ed altri, potrebbero

costituire target molecolari estremamente interessanti nell'ambito della chemioterapia antivirale.

La disponibilità di strategie sintetiche basate su una chimica semplice ed efficiente, quale la chimica combinatoriale e sull'utilizzo di materiali di partenza poco costosi e facilmente reperibili costituisce un obiettivo molto ambito per la sintesi di mimetici di composti naturali.

La ricerca attuale nel campo dei nucleosidi ad attività antivirale, ma anche antitumorale, antifungina, etc. è indirizzata verso la scoperta di nuovi analoghi caratterizzati da un'ampio spettro di attività biologiche e da un'elevata specificità d'azione al fine di minimizzare gli effetti collaterali di queste molecole. Per rispondere all'esigenza di avere, in tempi brevi, un elevato numero di analoghi nucleosidici da saggiare nelle loro eventuali proprietà farmacologiche, in questa tesi è stata messa a punto una strategia di sintesi in fase solida di una collezione di analoghi ciclici e aciclici (2',3'-seco-) dell'inosina, differenziati per la lunghezza della catena alchilica lineare in posizione N-1 della base ipoxantina. In letteratura sono riportati diversi esempi di strategie combinatoriali in fase solida per la preparazione di librerie di nucleosidi e piccoli oligonucleotidi analoghi. In ogni caso, qualunque sia l'approccio seguito, un requisito fondamentale è la scelta di un *linker* appropriato che funga da ponte tra il nucleoside e il supporto solido. Esso deve essere sufficientemente stabile alle condizioni di reazione usate durante la sintesi, ma allo stesso tempo abbastanza labile nelle condizioni di rimozione dal supporto solido senza inficiare l'integrità dei prodotti d'interesse.

La strategia seguita per la sintesi combinatoriale in fase solida di derivati N-1-alchilati dell'inosina è basata su studi<sup>[15]</sup>, condotti in precedenza dal gruppo di ricerca presso cui ho svolto il lavoro di ricerca, sulla reattività dell' N-1-dinitrofenil-2'-deossi-inosina nei confronti di nucleofili all'azoto, al fine di ottenere inosine N-1 sostituite e AICAR derivati. Questo approccio è stato utilizzato anche da altri gruppi di ricerca per introdurre una base purinica modificata all'interno di oligonucleotidi. Per estendere queste reazioni alla strategia combinatoriale in fase solida, si è usato un supporto polistirenico, disponibile commercialmente, a cui è legato covalentemente il gruppo MMTCl (monometossi-tritil-cloruro). Questo tipo di supporto, già utilizzato efficientemente, come riportato in letteratura, per legare nucleosidi o zuccheri, ha la peculiarità di essere stabile in condizioni alcaline, e viceversa molto labile in condizioni acide. A tale supporto è stato ancorato l'ossigeno 5' del ribosio della 2',3'-O-isopropiliden-N-1-(2,4-dinitrofenil)-inosina, con formazione del corrispondente legame etero 5'-O-tritilico.

Ancorato il nucleoside al supporto solido, è stata saggiata la reattività al carbonio C-2 dell'anello purinico nei confronti di diversi nucleofili amminici al fine di costruire una prima libreria di nucleosidi modificati. La procedura ha fornito, previo distacco dal supporto solido, i prodotti desiderati in ottima resa.

Passo successivo è stato quello di studiare la reattività dell'anello ribosidico modificando le posizioni 2' e 3'. Questo ha previsto l'aggancio diretto alla resina del nucleoside N-1-(2,4-dinitrofenil)-inosina non protetto alle funzioni ossidriliche 2' e 3', vista l'instabilità del legame 5'-O-etero alle condizioni di rimozione dell'isopropilidene. La condensazione tra il nucleoside e il supporto solido è risultata selettiva a carico delle funzioni alcoliche primarie, in accordo con quanto riportato in letteratura. Dopo alchilazione della posizione N-1 dell' inosina, seguendo la strategia prima descritta, è stata presa in esame la reazione di scissione del legame C-C 2'-3' da parte di composti ossidanti come il periodato ( $\text{IO}_4^-$ ). L'azione del periodato di sodio ( $\text{NaIO}_4$ ) sui nucleosidi deprotetti è ben conosciuta<sup>[16]</sup>. La



dialdeide risultante non è stata isolata, ma ridotta direttamente con sodio boroidruro ( $\text{NaBH}_4$ ). I prodotti di interesse sono stati ottenuti in buona resa.

Lo studio ha riguardato quindi la sintesi di nuovi nucleosidi analoghi quali potenziali agenti terapeutici ed ha condotto all'ottenimento di derivati che verranno poi sottoposti a saggi biologici per poterne testare l'eventuale attività terapeutica<sup>[17]</sup>.

- **Sintesi di nuovi derivati del cIDPR (ipoxantina-difosfo-ribosio ciclico) progettati come nuovi analoghi del cADPR, un potente secondo messaggero coinvolto nella mobilizzazione dello ione  $[\text{Ca}^{+2}]$ ;**

Nell'ambito di tale tematica mi sono interessata della sintesi di nuovi analoghi del cADPR, al fine di saggiarne le eventuali proprietà farmacologiche, di comprendere la relazione tra la struttura e l'attività di tale molecola, intimamente coinvolta nella regolazione della concentrazione di calcio intracellulare. Il cADPR esplica una funzione fondamentale nella regolazione del  $\text{Ca}^{+2}$ , pertanto risulta di grande interesse disporre di un numero maggiore di informazioni circa i meccanismi attraverso i quali quest'azione viene regolata. Purtroppo lo studio del ruolo fisiologico del cADPR è fortemente limitato dal ruolo transiente di tale molecola<sup>[18]</sup>. Allo scopo di ottenere analoghi del cADPR di pari attività, ma caratterizzati da un più stabile legame N-1 ribosidico, sono stati sintetizzati numerosi composti strutturalmente correlati ad esso, da utilizzarsi sia come modelli per comprendere la relazione struttura-attività, sia per le loro eventuali proprietà farmacologiche.

In letteratura sono riportati diversi esempi di analoghi stabili del cADPR, in alcuni casi anche più attivi dell'analogo naturale stesso. Uno dei primi è stato l' N-1-aristeromicina-adenina-difosforibosio ciclico<sup>[19]</sup> in cui l'ossigeno del ribosio in N-1 è sostituito da un gruppo metilenico. Questo composto risulta essere resistente sia all'idrolisi chimica, sia a quella enzimatica dal momento che non c'è più il legame N-1glicosidico presente nel cADPR. Alcune modifiche strutturali sull'adenina forniscono derivati più potenti del cADPR, è il caso del 3-deaza-cADPR avente un'attività 70 volte superiore<sup>[20]</sup>. Ulteriori modifiche strutturali includono il legame pirofosfato: il cATPR è, non solo più stabile, ma anche più attivo nell'indurre il rilascio di  $\text{Ca}^{+2}$  (20 volte maggiore) rispetto alla controparte cADPR<sup>[21]</sup>. Ulteriori studi struttura-attività, eseguiti utilizzando derivati opportunamente modificati al ribosio in posizione N-1 e alla funzione amminica in C-6 dell'adenina, indicano che questi non sono punti critici nell'attività del cADPR. Pertanto il cIDPR e diversi suoi analoghi sono stati sintetizzati. Essi mostrano possedere un'aumentata stabilità rispetto al cADPR e sembrano conservarne l'attività in alcuni sistemi cellulari, come le cellule-T, in cui sono stati saggiati<sup>[22]</sup>. Pochi studi sono stati condotti per provare la farmacoforicità del ribosio in posizione N-9 della base adenina. Preliminari studi suggeriscono che il gruppo 3'-ossidrilico è essenziale nell'attività di rilascio del  $[\text{Ca}^{+2}]$ . Risulta interessante, pertanto, investigare ulteriormente su tale residuo, allo scopo di verificare se l'attività del cADPR sia legata solo alla presenza dei gruppi OH sul ribosio in N-9 o necessiti anche della rigidità dell'unità ribosidica. La sintesi di nuovi analoghi N-9-aciclici che conservano solo le funzioni ossidriliche del ribosio ma non la rigidità dell'anello, è parte del mio lavoro di ricerca svolto nei tre anni di dottorato<sup>[23]</sup>.

- **Sintesi di TFO (oligonucleotidi formanti tripla elica) coniugati ad un residuo di acridina in prossimità del sito di inversione 3'-3', che, fungendo da**

**intercalante, stabilizzi la struttura triplex, compensando eventuali destabilizzazioni dovute alla giunzione 3'-3'.**

Per quanto concerne questo punto, l'unica metodica sintetica finora proposta risulta alquanto complessa e di non facile applicazione. Con questo lavoro si vuole proporre un'alternativa sintetica più semplice che consenta la realizzazione di ON-coniugati ad un residuo di cloro-metossi-acridina.

A tale scopo si è pensato di sfruttare la reattività all'alchilazione mostrata dall'N-3 di una timidina fiancheggiante il legame fosfodiesterico 3'-3' e la stabilità dimostrata dai complessi a doppia e tripla elica di DNA contenenti intercalanti, o altri gruppi legati all'N3 di un residuo di timidina<sup>[24]</sup>. Tra le varie molecole note per le loro proprietà intercalanti e stabilizzanti triplex, la scelta è ricaduta sull'acridina sia per l'elevato numero di derivati commercialmente disponibili, ma anche perché quando coniugata a sequenze ON è capace di indurre un più selettivo *cleavage* del DNA, in presenza di ioni metallici e per esposizione a fotoirradiazione. La formazione e la stabilità dei complessi a tripla elica, ottenuti usando come TFO l'ON-coniugato all'acridina, su descritto, sono state valutate a diversi pH, nell'intervallo 5.5-7.0, mediante studi di denaturazione termica UV e CD (dicroismo circolare), risonanza magnetica nucleare e studi di meccanica e dinamica molecolare<sup>[25]</sup>. La stabilità di tali complessi è stata confrontata con quella ottenuta utilizzando le corrispondenti sequenze oligonucleotidiche naturali<sup>[25]</sup>. Le informazioni provenienti da tali studi forniscono altresì valide indicazioni per la razionalizzazione delle modifiche chimiche più adatte per gli ON.

- **Sintesi e caratterizzazione di nuovi analoghi oligonucleotidici capaci di formare quadruple eliche di DNA monomolecolari, o "oligonucleotidi a grappolo".**

Le strutture a quadrupla elica sono caratterizzate dalla formazione di quartetti di guanine connesse da legami idrogeno di tipo Hoogsteen (quadruplex parallele) o reverse- Hoogsteen (quadruplex antiparallele). Questi complessi possono essere classificati in base al numero dei filamenti oligonucleotidici che li compongono (uno, due, quattro) e in base all'orientamento degli ON formanti il complesso (parallelo o antiparallelo). Benché la maggior parte delle strutture G-quadruplex con attività biologica siano di tipo antiparallelo, recenti studi hanno evidenziato interessanti attività anche per quadruplex parallele. Infatti quadruplex parallele interagiscono selettivamente con alcune proteine<sup>[26]</sup>, inoltre, è stata osservata per alcuni di questi complessi, una attività anti-HIV<sup>[5]</sup> ed anche attività di aptamero verso l'ematoporfirina<sup>[5]</sup>. Quadruplex parallele si formano quasi esclusivamente in complessi tetramolecolari. Sfortunatamente, in vivo, la formazione intermolecolare di complessi quadruplex è molto lenta e necessita di alte concentrazioni di ON, non facilmente raggiungibili in condizioni fisiologiche. Gli sfavorevoli parametri cinetici e termodinamici sono quindi un ostacolo alla potenziale utilizzazione farmacologica di questo tipo di complessi. Di qui l'interesse per la sintesi e la caratterizzazione strutturale di un nuovo tipo di quadruplex parallela monomolecolare formata da oligonucleotidi legati a "grappolo"<sup>[27]</sup>. In particolare, la struttura oligonucleotidica a grappolo prevede che i quattro filamenti, formanti il complesso quadruplex, siano legati reciprocamente per le estremità 3' attraverso una molecola ramificata simmetricamente. Questa nuova tipologia di quadruplex dovrebbe essere caratterizzata da una maggiore stabilità e da una veloce cinetica di strutturazione

rispetto ai corrispondenti complessi tetramolecolari, ipotesi confermate dai risultati. La sintesi della molecola "linker" tetraramificata (catene dialchilfosfato) avviene su supporto polimerico insolubile. Ogni ramificazione contiene una funzione alcolica terminale a cui è possibile legare la catena oligonucleotidica. Studi iniziali sono stati effettuati sintetizzando ON a grappolo aventi sequenza TGGGT. Modulando opportunamente i gruppi protettori delle funzioni alcoliche della struttura ramificante e la tipologia di accrescimento della catena oligonucleotidica, sono state ottenuti ON a grappolo formanti strutture quadruplex con orientamento dei filamenti predefiniti (paralleli, antiparalleli o con orientazione mista) [28-29]. La maggiore stabilità ipotizzata per questa nuova classe di analoghi è confermata mediante esperimenti NMR e CD (dicroismo circolare), ha condotto ad estendere lo studio anche relativamente a sequenze oligonucleotidiche naturali formanti strutture quadruplex caratterizzate da bassa stabilità, ad esempio contenenti A-tetrad e T-tetrad<sup>[30-31]</sup>. Questa nuova tipologia di quadruplex sarà inoltre estesa a complessi contenenti filamenti ibridi DNA-PNA. L'accresciuta stabilità delle quadruplex monomolecolari potrebbe portare alla formazione di stabili quartetti di guanine nei tratti di PNA, finora mai osservati.

Il mio progetto di ricerca, in tale ambito, ha previsto, inoltre, l'ottimizzazione della lunghezza del linker tetraramificato ai fini sia di una maggiore stabilizzazione della struttura quadruplex, sia per poter selezionare quadruplex antiparallele da strutture sintetizzate a filamenti pre-orientati a polarità inversa<sup>[32]</sup>. È noto che la polarità dei filamenti in una quadruplex riveste fondamentale importanza nelle interazioni con proteine o altre molecole. Una quadruplex può essere parallela o antiparallela in ragione della sua molecolarità, della sequenza di basi e, in alcuni casi, in ragione della sequenza e tipologia dei *loop* di connessione che si realizzano tra i filamenti formanti la quadruplex. La strategia degli ON a grappolo trascende da tutte queste limitazioni e consente di preparare quadruplex antiparallele, purché si utilizzino linker di opportuna lunghezza. L'eccessiva lunghezza del linker tetraramificato, infatti, sembra compromettere la formazione della quadruplex antiparallela favorendo la formazione della struttura parallela per ripiegamento dei *loop* alchilici, come dimostrano i dati ottenuti da calcoli di meccanica molecolare condotti in tale ambito. La lunghezza del linker sembra anche giocare un ruolo di cruciale importanza sulla riproducibilità e velocità dei meccanismi di melting e refolding delle strutture quadruplex a grappolo. Preliminari studi CD/T (dicroismo circolare al variare della temperatura), eseguiti su vari analoghi caratterizzati da una diversa lunghezza del linker, indicano che la specie con l'unità tetraramificante più corta sia il deputato migliore su cui fare studi su nanomotori. Infine una particolare attenzione è stata rivolta alle indagini strutturali sui complessi formati dagli oligonucleotidi sintetizzati. I dati ottenuti mediante impiego di tecniche spettroscopiche, quali CD, NMR mono- e bidimensionale, eseguiti a diverse temperature ed in vari tamponi, hanno consentito di caratterizzare le strutture di tali complessi.

È stato inoltre valutato anche il diverso effetto stabilizzante dovuto alla presenza di cationi potassio e sodio. I dati ricavati dalla spettroscopia NMR, inseriti in metodi di calcolo (Cyana e Dyana) consentiranno di proporre valide strutture per i complessi studiati.

## Riferimenti Bibliografici

1. Mazza, P.; *Pharm. Chem.Bull.*, **1992**, 131, 355-362
2. Knauert,M.P.; Blazer,P.M.; *Human Molecular Genetics*, **2001**, 10, 2243-2251
3. Glosal, G.; Muniyappa,K.; *Biochemical and Biophysical Research Comm.*, **2006**, 343, 1-7
4. Beal, P.A.; Dervan, P.B.; *J.Am.Chem.Soc.*, **1992**, 114, 4976-4982
5. a) Arthanari, H.; Boltono, P.; *Chemistry & Biology*, **2001**, 8, 221-230  
 b) Zhang, N.; Phan, A.T.; *J.Am.Chem.Soc.*, **2005**, 127, 17277-17285  
 c) Phan, A.T.; Patel, D.J.; *Nature.Chem.Biol.*, **2005**,1, 167-173  
 d) Phan, A.T.; Patel, D.J.; *Proc.Natl.Acad.Scis.USA*; **2005**, 102, 634-639
6. Jing, N.; Marchand, C.; *DNA Cell Biol.*, **2001**,20, 499-508
7. Pileur, F.; Andreola, M.L.; *Nucleic.Acids Res.*,**2003**, 31, 5776-5788
8. Greider, C.W.; *Proc.Natl.Acad. Sci.USA.*, **1998**, 95, 90-92
9. Kerwin, S.M.; *Curr.Pharm.Design* ,**2000**, 6, 441-471.
10. Siddiqui-Jain, A.; Grand, C.; *PNAS*, **2002**, 99, 11593-11598
11. Delaney, S., et al; *Biochemistry*, **2003**, 42, 14159–14165
12. de Champdoré, M., et al , *Chem. Commun.*, **2004**,1756–1757
13. Guse, A.H. ; *Curr.Mol.Med.*, **2004**, 4, 239-248
14. Shuto,S.; Shirato, M. ; *J.Org.Chem.*,**1998**, 63, 1986-1994
15. De Napoli L.; Messere A.; *J. Chem. Soc.Perkin Trans.1*, **1997**, 2079-2082
16. Chamakura V.; Varaprasad, Qazi Habib; *Tetrahedron*, **2003**, 2297-2307
17. Borbone, N.; Amato, J.; et al.; *Nucleosides Nucleotides and Nucleic Acids*, accettato per la pubblicazione
18. Hutchinson, E.J.;Taylor, B.F.; *J.Chem.Soc.Comm.*,**1997**,19, 1859-1860
19. Shuto,S.; Fukuoka, M.; *J.Am.Chem.Soc.*, **2001**, 123,8750-8759
20. Wang,L.; Aarhus,R.; *Biochem.. Biophys.Acta*, **1999**, 1472, 555-564
21. Aarhus,R.; Gee,K.; *J.Biol.Chem.*, **1995**, 270, 7745-7749
22. Gu, X.; Yang, Z.; *J.Med.Chem.*, **2004**, 47, 5674-5682
23. Amato, J.; Oliviero, G.; *Nucleosides Nucleotides and Nucleic Acids*, **2005**, 24, 735-738
24. Amato, J.; Oliviero, G.; *Nucleosides Nucleotides and Nucleic Acids*, accettato per la pubblicazione
25. Kubuota, M.; Ono, K.; *Nucleic.Acids Res.*,**2002**, 2, 33-34
26. Amato, J.; Galeone, A.et al.; *Eur.J.Org.Chem.*, **2004**, 2331-2336
27. Borbone,N.; Oliviero, G. et al; *Tetrahedron Lett.*, **2004**, 45, 4869-4872
28. Amato, J.; Borbone, N., et al. *Nucleosides Nucleotides and Nucleic Acids*, **2005**, 24, 739-741
29. Amato, J.; Borbone, N., et al.; *Bioconjugate Chemistry*, **2006**, 17, 889-898
30. Amato, J.; Borbone, N., et al; *Nucleosides Nucleotides and Nucleic Acids*, **2005**, 24, 443- 446
31. Amato, J.;Oliviero, G.,et al; *Biopolymers*, **2006**, 81, 194-201
32. Amato, J; Borbone, N., et al; *Nucleosides Nucleotides and Nucleic Acids*, accettato per la pubblicazione

## Prefaction

In these researches, carried out in three years of PhD in Biotechnology, (Molecular-Biotechnology), I focused my attention on Nucleic Acid chemistry. Particularly, I investigated two main aspects of this field: the chemistry and the structural characterization of some nucleic acid derivatives and the synthesis of several new nucleoside and nucleotide analogues.

For a greater convenience of exposition, I wish to report the results arranged in two main sessions:

- Synthesis and structural characterization of oligonucleotides (ON) analogues.
- Synthesis of nucleosides and nucleotides analogues.

To these researches the publication of the following articles ensues:

1. Amato J., Galeone A., Oliviero G., Mayol L., Piccialli G., Varra M. **Synthesis of 3'-3'-linked pyrimidine oligonucleotides containing an acridine moiety for alternate strand triple helix formation.** European Journal of Organic Chemistry (2004), (11), 2331-2336.
2. Oliviero G., Amato J., Varra M., Piccialli G., Mayol L. **Synthesis of a new N-9 ribityl analogue of cyclic inosine diphosphate ribose (cIDPR) as a mimic of cyclic ADP ribose (cADPR).** Nucleosides, Nucleotides & Nucleic Acids (2005), 24(5-7), 735-738.
3. Oliviero G., Amato J., Borbone N., Galeone A., Varra M., Piccialli G., Mayol, L. **Unusual monomolecular DNA quadruplex structures using bunch-oligonucleotides.** Nucleosides, Nucleotides & Nucleic Acids (2005), 24(5-7), 739-741.
4. Oliviero G., Amato J., Borbone N., Galeone A., Varra M., Piccialli G., Mayol L. **Synthesis and characterization of DNA quadruplexes containing T-tetrads formed by bunch-oligonucleotides.** Biopolymers (2006), 81(3), 194-201.
5. Oliviero G., Amato J., Borbone N., Galeone A., Petraccone L., Varra M., Piccialli G., Mayol L. **Synthesis and Characterization of Monomolecular DNA G-Quadruplexes Formed by Tetra-End-Linked Oligonucleotides.** Bioconjugate Chemistry (2006), 17(4), 889-898.
6. Oliviero G., Borbone N., Amato J., D'Errico S., Piccialli G., Varra M. and Mayol L. **Synthesis of New Ribose Modified Analogues of Cyclic Inosine bis-phosphate Ribose (cIDPR),** in press for Nucleosides, Nucleotides & Nucleic Acids.
7. Borbone N., Mayol L., Miccio L., Oliviero G., Amato J., Pesce G., Piccialli G. and Sasso A.. **Optical Tweezers as a Probe for Oligodeoxyribonucleotides Structuration,** in press for Nucleosides, Nucleotides & Nucleic Acids.
8. Borbone N., Oliviero G., Amato J., D'Errico S., Galeone A., Piccialli G., and Mayol L. **Synthesis and Characterization of Tetra-End Linked Oligonucleotides capable of forming Monomolecular G-Quadruplexes,** in press for Nucleosides, Nucleotides & Nucleic Acids.

9. Oliviero G., Amato J., D'Errico S., Borbone N., Piccialli G. and Mayol L. **Solid phase synthesis of nucleobase and ribose modified inosine nucleoside analogues**, in press for Nucleosides, Nucleotides & Nucleic Acids.
10. Oliviero G., Amato J., Borbone N., D'Errico S., Piccialli G. and Mayol L. **Synthesis of N-1 and ribose modified inosine analogues on solid support**, in press for Tetrahedron Letters.

# Chapter 1

## *“General Introduction”*

### **1.1. *Synthesis and structural characterization of oligonucleotide (ON) analogues***

In recent years, the most important goal of researches in the field of medicinal chemistry is the achievement of a drug capable to interact with its target in a specific or, at least, highly selective way. The rational design of such a drug requires a detailed knowledge of its target, which is not always available if this is a protein. More favourable is certainly the case where the target is a DNA or a RNA, since structures and functions of these classes of molecules are better known and, above all, mechanisms of molecular recognition in which they are involved have been clarified in details. In order that a molecule, as a potential drug, may specifically interact with a particular nucleic acid, it has to be able to read a sequence not shorter than 16-18 bases. At the best of the present knowledge, the only molecule capable to do this is another nucleic acid. This is the reason why, in the last decade, interest towards synthetic oligonucleotides (ONs) has been growing, in consideration of their use as selective regulators of genetic expression. Oligonucleotides, addressed to the control of a specific gene expression, can be catalogued on the base of their target and mechanism of action, in “antisense” or “antigene”, when they target a specific mRNA or a tract of duplex DNA, respectively, or as “aptamers”, when selectively recognize a specific protein. Antisense strategy, for gene regulation, is based on the specific recognition of the target mRNA sequence by synthetic ONs having complementary sequence by formation of a duplex structure stabilized by Watson-Crick hydrogen bonds. In this way the formation of duplex DNA-RNA hybrids is achieved, which can give rise to various processes through which the translation of the message is efficiently inhibited. In antigene strategy, an ON acts at a higher level, having as a target a double-stranded DNA, with which it forms a triple helix (triplex) blocking its transcription. The recognition of a double stranded DNA by an ON is highly specific, being based on the formation of base triplets stabilized by hydrogen bonds, according to Hoogsteen or reverse Hoogsteen schemes. An ON represents an ideal drug from a pharmacodynamic point of view, being capable to act in an absolutely specific way and at low cellular concentration. On the other hand, an ON, as a drug, is characterized by a number of pharmacokinetic drawbacks. Potential therapeutic applications of natural ONs are, in fact, limited by the poor cellular uptake and the rapid degradation promoted by cellular nucleases. To overcome these problems, several ON analogues have been designed and synthesized and, after more than two decades of studies, the first antisense ONs have been recently introduced in therapy as antiviral drugs, while others are currently in clinical trials as antitumor agents. Antigene applications, in addition to showing the same problems as the antisense strategy, are further limited by the request of homopurine-homopyrimidine tracts of 16-18 bases on the target, namely a double-stranded DNA. To expand the number of biologically relevant DNA sequences, as targets by triplex structures formation, Horne and Dervan, then followed by other research groups, proposed the use of ONs containing an inversion of polarity motif, i.e. a 3'-3' internucleosidic linkage, to target mixed sequence duplexes of the type (purine)*m*(pyrimidine)*n*. In order to stabilize the 3'-3' internucleosidic linkage, my research work was addressed into synthesis of oligonucleotides with a 3'-3' inversion

of polarity, containing an acridine moiety attached to the nucleotide base flanking the 3'-3' phosphodiester bond, able to form more stable alternate strand triple helix complexes. These have been investigated by UV melting studies and CD experiments.

A further therapeutical application of synthetic ONs is their potential use as aptamers. An aptamer is an ON able to adopt a proper conformation allowing it to specifically interact with a particular protein, thus modulating its function. The evidence that short DNA and RNA fragments are able to specifically bind several proteins lead to the consequence that they are no longer considered as only passive carriers of genetic information, but also as bio-molecules involved in many key cellular processes. ONs as aptamers are single, double or quadruple strand nucleic acids able to selectively recognize, bind and inhibit specific proteins involved in gene control expression or in the development of particular disease. ONs characterized by a G-rich sequence and adopting a G-quadruplex structure, could be considered useful aptamers for the control of the cellular senescence and carcinogenesis. Furthermore, recent studies on biological activities of synthetic ONs forming quadruplex structures, suggest that these molecules can act as inhibitors of important proteins such as thrombine and HIV integrase. However, the biological importance of these complexes lies in their possible presence in gene telomeric sequences. Quadruplex structures (G-quadruplex), generated by a DNA tract containing a G-rich sequence, are located at the 3'-end of telomeres, where they probably play an important role in the regulation of the telomerase activity. Telomerase is the most important enzyme involved in telomere maintenance in tumour cells, identified so far. The high activity of this enzyme inhibits the senescence and the death of tumour cells. This hypothesis has originated a large number of studies on the inhibition of the telomerase by ONs forming quadruplex structures (G-quadruplex-aptamers). Quadruplex complexes are characterized by the formation of G-quartets connected by Hoogsteen hydrogen bonds. G-quadruplexes can be classified depending on the number of self-associating strands that form the structure (one, two or four strands) and are further distinguished by the orientation of the strands (parallel or antiparallel). Even though most of the G-quadruplex structures involved in biomolecular events belong to the antiparallel type, yet evidences have been provided that some proteins are able to interact with parallel G-quadruplex as well. Biological relevance of such structures has further risen due to the discovery of ONs forming parallel quadruplexes, including aptamers characterized by an anti-HIV activity and hematoporphyrine binding aptamers. Parallel quadruplex are found almost exclusively in tetramolecular complexes. Unfortunately, the in vivo formation of intermolecular quadruplexes is very slow, requiring high ONs concentration. These unfavourable kinetic and thermodynamic parameters could be disadvantageous in view of their potential therapeutic use. In this contest, another goal of my research program has been aimed to the synthesis and the structural characterization of a new kind of monomolecular G-quadruplex formed by ONs linked through a tetra-end-linked (TEL) structure. Particularly the structural feature of this TEL-ON, forming the quadruplex, is the presence of four ON strands whose 3'-ends are linked together by a symmetric tetra-branched spacer. The formation of a monomolecular quadruplex could be characterized by favourable kinetic and thermodynamic parameters and the complex could be more stable than the corresponding tetramolecular counterpart. Preliminary results, on the synthesis and stability of this new kind of quadruplexes, carried out in our laboratory, confirmed this hypothesis. The synthesis of the symmetric tetra-branched linker have been carried out on polymeric support. This branched structure bears four primary alcoholic functions prone to ON chain assembly (for example TGGGT o TGGXGGT, where X is a variable base). By choosing appropriate alcoholic protecting groups of the linker structure and the ON chain assembly procedure, TEL-ONs having predetermined strand orientation (parallel, antiparallel or both)



or containing modified bases, have been obtained. Moreover, quadruplex complexes, containing less stable quartets such as A-tetrads, C-tetrads, T-tetrads, mixed tetrads, as well as tetrads formed by modified nucleosides could be synthesized. In this context, the chemical synthesis of TEL-oligonucleotides prone to form quadruplex structures containing G- and T-tetrads has been performed. This new kind of stable quadruplex structure could be extended to the complexes containing DNA-PNA hybrid strands. It is reasonable to hypothesize that the increased stability of these new monomolecular quadruplexes could permit the formation of PNA guanine quartets, not observed so far. Pharmacological structural investigations on complexes formed by synthesized ONs and synthetic models of their natural targets (mRNA, DNA or proteins), have been essentially carried out by NMR spectroscopy associated with mechanic and molecular dynamics calculations, in order to allow a better comprehension of specific recognition mechanisms. This is useful to design new rational structural modifications to be introduced in synthetic ONs, to optimize their properties and the related biological response.

Briefly, the first session of my thesis research have been employed and addressed on :

- *Synthesis of TFO (Triplex Forming Oligonucleotides) containing a 3'-3' polarity inversion site conjugated to moieties (acridine residue) capable to stabilize the triple helix complexes or to supply particular biological activities.*
- *Synthesis and characterization of new kind ON analogues able to form stable monomolecular quadruple helices DNA complexes (Tetra-End-Linked Oligonucleotides TEL-ON).*  
*In this context, studies on the length of tetra-end-linked linker on the orientation of the ON strands aimed at obtaining stable quadruplex structures and studies on quadruplex complexes, containing less stable quartets such as T-tetrads, have been performed, too.*

Then, special attention has been devoted to structural investigation on the complexes formed by the synthesized oligonucleotides. The use of spectroscopy methods, nuclear magnetic resonance (NMR) and mass spectrometry (MS) above all, associated with molecular mechanism and dynamic studies, result fundamental to allow a deep insight into molecular recognition mechanism in which the synthesized molecules are involved. Particularly, all synthesized molecules have been purified by HPLC and structurally characterized by MS spectrometry, NMR spectroscopy and CD analysis. The stability of triple helix complexes have been evaluated by thermal denaturation studies using both UV and circular dichroism spectroscopy (at several pH and buffer composition). Furthermore, the stabilization effect of Na, K cations for quadruplex complexes, have been evaluated. The obtained data have been allowed a further insight into specific recognition mechanisms. Physico-chemical studies on the stability of triple and quadruple helices have been also carried out. These studies provide detailed information on the energetic of DNA high-order structure formation on changing the pH, temperature and base composition. Differential scanning calorimetry (DSC), circular dichroism (CD) and computational methods have been used to characterize triple and quadruple helices formation. Finally, studies are currently in progress in order: i) to extend the synthetic strategy employed to link other agents near to 3'-3' phosphodiester linkage, thus furnishing a convenient method to stabilize alternate strand triple helices; ii) to assay TEL-ON in their potential wide range of biological properties.

## 1.2. *Synthesis of nucleoside and nucleotide analogues*

The second part of my research thesis has been aimed to the synthesis of new nucleosides and nucleotides analogues.

Recently, medicinal chemistry research emphasizes the need of nucleosides and nucleotides analogues in high number and diversity, to be used as building blocks in the synthesis of new ON generations. Apart from being the genomic building blocks, nucleosides interact with roughly one-third of protein classes in the human genome, including not only polymerases but also kinases, reductases, motor proteins, membrane receptors and structural proteins, all targets of therapeutic importance in biological systems. Naturally occurring nucleosides analogues demonstrate selective activities, such as protein synthesis inhibitors (puromycin), glycosyl transferase inhibitors (tunicamycin), and methyl transferase inhibitors (sinefungin). Synthetic analogues of nucleosides have been the cornerstone of antiviral therapy over the past 30 years. Actually, approximatively 50 nucleosides and nucleotide-related drugs have been pushed to the market and over 80 nucleosides/tides are in the (pre) clinical studies for various therapeutic indications, including not only antivirals but also antineoplastics, antifungins and others. However, some drugs such as AZT, ddI, 3TC, etc. rapidly develop drug resistance and show mitochondrial, bone-marrow and other toxicity. Therefore, it is essential to discover novel nucleoside drugs for critical medical needs. The availability of high-throughput screening capabilities, together with the combinatorial synthesis of small organic molecule libraries, offer a unique opportunity to accelerate the discovery in these fields. At this purpose, recently, I'm looking for the solid phase synthesis of nucleoside and nucleotide libraries. Preliminary studies in this contest allowed the realization of a small library of N1-alkylated inosine, N1-alkylated- 2',3'-seco-inosine and AICAR derivatives. Finally, nucleosides and nucleotides are also rated as the central molecules of metabolism, playing a strategic role as endogenous regulators of cell functions. An example is represented by cyclic ADP-ribose, a 18-membered cycling nucleotide. Cyclic ADP-ribose (cADPR) is an endogenous metabolite shown to mobilize intracellular  $\text{Ca}^{+2}$  in various cells, indicating that it is a general mediator involved in  $\text{Ca}^{+2}$  signalling. It is also known to mobilize intracellular  $\text{Ca}^{+2}$  more actively than inositol-1,4,5-triphosphate ( $\text{IP}_3$ ) by a mechanism completely independent of  $\text{IP}_3$ . Since  $\text{Ca}^{+2}$  are regulators of diverse cell functions, e.g. muscle contraction, secretion of neurotransmitters, hormones and enzymes fertilization of ovocyte and lymphocyte activation and proliferation, the cADPR signalling pathway may become a valuable target for pharmaceutical intervention. The usefulness of cADPR is not only as a pharmacological tool, but also in the elucidation of the mechanism of cADPR action, labelling of binding proteins or antagonizing  $\text{Ca}^{2+}$  mobilization. cADPR is synthesized from  $\text{NAD}^+$  by ADP-ribosyl-cyclase in cells, and it is characterized by a very labile N1-glycosidic bond, which is rapidly hydrolyzed both enzymatically, by cADP-hydrolase, than non-enzymatically, to give ADP-ribose even in a neutral aqueous solution. Therefore, stable analogues of cADPR that have a  $\text{Ca}^{+2}$  -mobilizing activity in cells similar to that of cADPR are urgently required. In this contest my research work have been addressed on synthesis of new stable analogues of cADPR, particularly N-9 ribose modified derivatives.

Briefly, my studies in nucleoside and nucleotide analogues research have been addressed on such targets:

- *Synthesis of combinatorial libraries of nucleosides, specifically inosine analogues.*
- *Synthesis of new cIDPR (cyclic hypoxanthine-diphospho-ribose) analogues, designed as novel stable mimics of cADPR, a potent  $\text{Ca}^{2+}$ -mobilizing second messenger.*

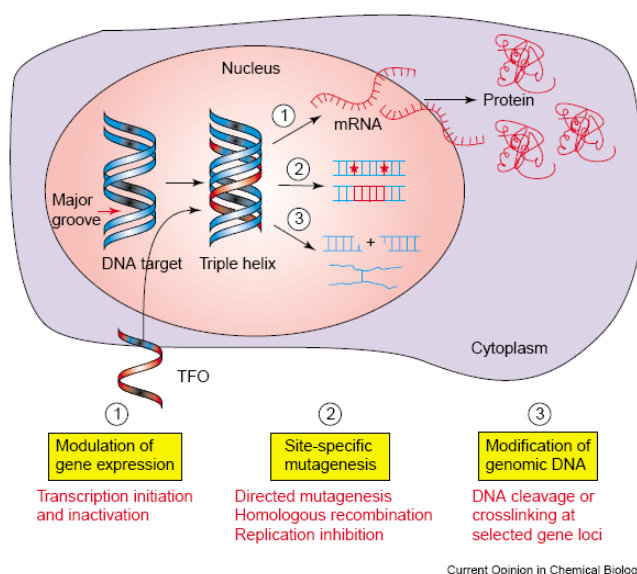
Studies are currently in progress in order to obtain both new and largest libraries of nucleoside analogues and new cADPR analogues, that will be screened against a wide range of their respective biological assays.

## Chapter 2

***“Synthesis of TFO (Triplex Forming Oligonucleotides) containing a 3'-3' polarity inversion site conjugated to moieties (acridine residue) capable to stabilize the triple helix complexes or to supply particular biological activities”***

### Introduction

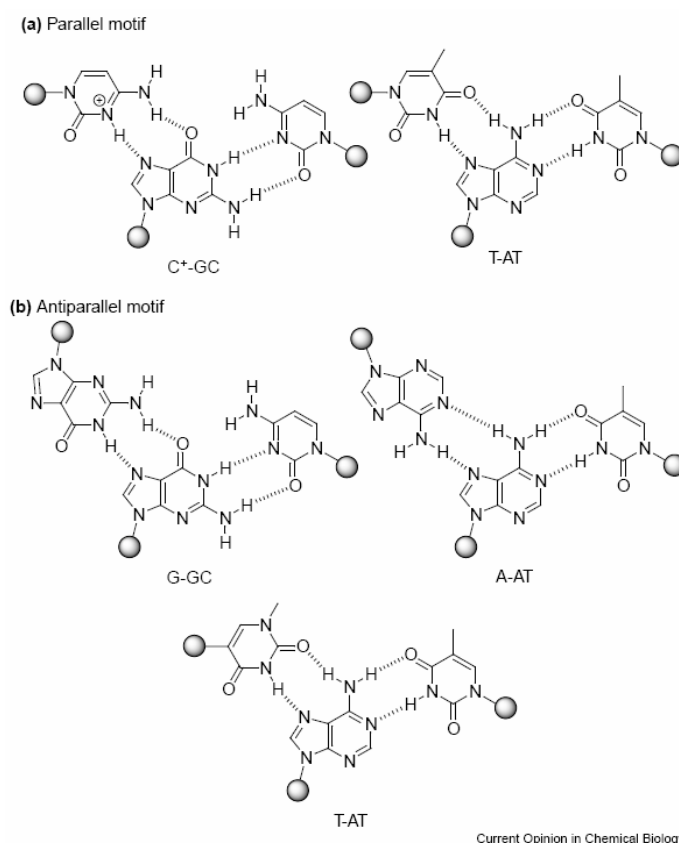
The remarkable conformational flexibility of nucleic acids allows for the formation of a great variety of structures besides the canonic B-DNA duplex. Among these structures, those involving more than two strands have received considerable attention in the last decade<sup>[1]</sup>. There are two main reasons for the intense interest in such multistranded structures. First, there is more than circumstantial evidence for their biological role in replication, transcription and recombination<sup>[2]</sup>. Second, several multistranded motifs are used as the basis of some therapeutic approaches<sup>[3]</sup>. Nucleic acids have the potential to form not only double helix structures but also triple and quadruplex structures<sup>[4]</sup>. Despite a lack of direct evidence for the tertiary structure of nucleic acids *in vivo*, the presence of H-DNA was often considered to regulate gene transcription and the sequence-specific DNA ligands of the triplex forming oligonucleotides (TFOs) have been proven to target specific mutations in somatic cells of adult mice after intraperitoneal injection<sup>[5]</sup>. TFOs bind specifically to duplex DNA and provide a strategy for site-directed modification of genomic DNA, figure 2.1. The recognition of double-stranded DNA by a single stranded TFO provides a technique for site-specific genome modification in living cells. Moreover, the oligonucleotide-directed triplexes were extensively applied for sequence-specific modification to recognize the double-stranded nucleic acids targets or the cleavage of nucleic acids<sup>[6]</sup>. These potential gene-targeted therapeutic and biotechnological applications have provoked considerable interest in triplex formation and stability.



**Figure 2.1:** Basic antigenic concepts and corresponding applications.

Fensefeld et al.<sup>[7]</sup> described the first example of the formation of triple helical nucleic acids structures in 1957. A decade later, transcription inhibition of *Escherichia coli* RNA polymerase by an RNA third strand was reported by Morgan and Wells<sup>[8]</sup>, demonstrating a potential biological role of triplex structures. Subsequently, synthetic oligonucleotides

designed to form triplexes have been used in a variety of applications to manipulate genes and gene function both in vitro and in vivo. TFOs have been used to inhibit protein binding to DNA, to inhibit gene expression, to inhibit DNA replication, to direct site-specific damage, to enhance recombination and to induce mutagenesis<sup>[9]</sup>. The process of down-regulating gene expression, through triple helix formation, is referred to as the “antigene strategy”. Triple helix formation has also been utilized in the development of artificial nucleases, created by tethering a cleavage-agent to a triplex forming oligonucleotide. Binding characteristics of TFOs depend on their nucleotide composition. Thus, oligonucleotides composed of pyrimidine bases (C and T) bind through Hoogsteen hydrogen bonds and are oriented parallel to the purine-rich strand of the target duplex, forming  $C^{+}\cdot G\cdot C$  and  $T\cdot A\cdot T$  triplets, figure 2.2.



**Figure 2.2:** Molecular structures of the canonical Hoogsteen and reverse Hoogsteen base triplets in the known parallel and antiparallel triple helical binding motifs.

TFOs composed of purine bases (G and A), or mixed purine/pyrimidine (G and T) bind in anti-parallel orientation through reverse Hoogsteen bonds, forming  $G^{+}\cdot G\cdot C$  and  $A^{+}\cdot A\cdot T$  or  $T^{+}\cdot A\cdot T$  triplets<sup>[10]</sup>. In certain circumstances GT-TFOs can bind in parallel orientation<sup>[11]</sup>. Protonation of cytosine is required to form two Hoogsteen bonds in  $C^{+}\cdot G\cdot C$  triplets, so that binding of CT-TFOs is favoured by low pH. Triplex formation by GA and GT-TFOs is pH independent. Binding, in each of these motifs, can occur in physiological conditions, although cytosines must be modified to overcome pH dependence of CT-TFOs. Purine-rich tracts are frequently found in gene promoter regions and TFOs, directed to these regulatory sites, have been shown to selectively reduce transcription of the targeted genes, likely by blocking binding of transcriptional activators and/or formation of initiation complexes. Triplex-mediated modulation has potential application in therapy since it can be used, for example, to reduce levels of proteins thought to be important in disease processes. TFOs

can also be useful molecular tools for studying gene expression and function. An example is the down-regulating expression of the c-myc oncogene in cancer cells. c-myc is an attractive target for antigene agents in cancer cells because its expression drives cell proliferation<sup>[12]</sup>. The importance of c-myc in cancer cell growth is emphasized by the findings that it is frequently amplified or involved in chromosomal rearrangements and that its expression is deregulated and augmented through various mechanisms in many cancers. Functions of the c-myc protein, in tandem with its obligatory binding partner Max, include transcriptional activation and repression. Both activities appear to promote cell proliferation. Recent studies demonstrated that reduced expression of inducible c-myc was sufficient to cause regression of hematologic tumours in mice, suggesting that this approach may have therapeutic potential. Several sequences, suitable for triplex formation, are present in the c-myc gene. Of particular interest is a highly conserved polypurine/polypirimidine tract in the c-myc P2 promoter region. P2 is the major c-myc promoter and gives rise to 75-90% of transcripts in almost all cells, with P1 contributing most of the remainder. TFOs directed to P2 sequence had transcriptional inhibitory activity in vitro and in cells.

## **2.1. DNA triple helix formation.**

The primary requirement for triple helix formation is a homopurine-homopyrimidine sequence in the target duplex. Typically, the third strand sequence is also homopurine or homopyrimidine. Depending on the nature of the third strand, there are two many categories of triple helices: pyrimidine\*purine:pyrimidine (pyr\*pur:pyr) or purine\*purine:pyrimidine (pur\*pur:pyr), where \* represents Hoogsteen base pair formation between the third strand and the purine strand of the duplex and : denotes the Watson-Crick base pair of the duplex. Pyr\*pur:pyr triplex is less stable at low pH because the cytosines, on the third strand, require protonation in order to form Hoogsteen hydrogen bonds with the purine strand of the duplex. In contrast, pur\*pur:pyr triplexes are stable at neutral pH. Thermal denaturation studies on pur\*pur:pyr triplexes often reveal a single transition, suggesting the simultaneous dissociation of all three strands. In most examples of a pyr\*pur:pyr triplexes, a biphasic transition is observed, with the third strand dissociating at a lower temperature than the duplex. This suggest that a pyrimidine third strand may be less stable, in general, than a purine third strand. Unlike the Watson-Crick base-paired double helix, in which the two strands are bound antiparallel to each other, the third strand of a triple helix can be either parallel or antiparallel with respect to the purine strand to which it is hydrogen-bonded. The polarity of the third strand in the triplex is dependent on the sequence as well as the base composition. A homopyrimidine third strand, containing cytosine and/or thymine, binds to the purine strand in a parallel orientation. The binding of a homopurine oligonucleotide to the purine strand depends on the sequence. Experimental evidence has been reported for both parallel and antiparallel orientations of the third strand in triplexes composed solely of G\*G:C triplets. But if the third strand contains both Gs and As, the polarity of the third strand will be antiparallel with respect to the purine strand, since As with anti-glycosidic bond conformations can form only reverse-Hoogsteen hydrogen bonds and hence direct the oligonucleotide in an antiparallel orientation<sup>[13]</sup>. Hogan and co-workers<sup>[14]</sup> have been reported antigene activity of oligonucleotides composed of G and T bases. However, there have been only a limited number of physical studies on triplexes formed by such sequences. Helene and co-workers<sup>[15]</sup> have demonstrated that the orientation of a GT-third strand depends on its sequence. Calculations indicate that a 10-mer triplex whose third strand is composed of equal number of G and T bases, the third strand will be antiparallel to the duplex purine

strand if there are three or more GpT or TpG steps; otherwise it will be parallel. According to the calculations of Sun and Helene, this can be explained in terms of a balance between the preference for the parallel orientation (Hoogsteen hydrogen bonding) in an all G-third strand and the better tolerance for the lack of isomorphism between G\*G·C and T\*A·T triplets in the antiparallel orientation (reverse Hoogsteen hydrogen bonding). Radhakrishnan et al.<sup>[16]</sup> have been carried out NMR studies on an intramolecular triplex composed of a third strand containing Gs and one T. These studies provided clear evidence for the distortion induced by insertion of a T\*A·T triplet within a stack of G\*G·C triplets.

### 2.1.1. *Hoogsteen base-pairing.*

In double stranded DNA, whether in A-, B-, Z-form, the complementary strands are held together solely through Watson-Crick (W-C) base-pairing geometry. The hydrogen bonds orchestrate base-pairing between dA·dT or dG·dC and maintain the geometric structure of regular double-helical DNA. Over 40 years ago, K. Hoogsteen discovered a novel type of hydrogen bonding between adenine and thymidine, termed Hoogsteen base-pairing geometry<sup>[17]</sup>. These pairings involve either purines or pyrimidines interacting with the sites on purine bases, that are not involved in W-C hydrogen bonding (N7 and O6 for guanine, N7 and N6 for adenine)<sup>[18]</sup>. Hoogsteen base pairs have quite different properties from W-C base pairs. The angle between the two glycosylic bonds (ca. 80° in the A.T pair) is larger and the C1'-C1' distance (8.6 Å) is smaller than in the regular geometry. In some cases, called reversed Hoogsteen base pairs, one base is rotated 180° with respect to the other<sup>[19]</sup>, figure 2.2. Since the discovery of Hoogsteen base-pairing, forms of DNA double helix containing both inter- as well as intra-strand Hoogsteen and reverse Hoogsteen base pairs have been discovered in crystal structures of undistorted B-DNA<sup>[20]</sup>, DNA complexed with intercalating drugs<sup>[21]</sup>, protein/DNA complexes<sup>[22]</sup> and RNA<sup>[23]</sup>. The Hoogsteen base pairs are frequently engaged by proteins and ligands to interact with DNA. With the recent discovery of the formation of Hoogsteen base pairs, during DNA replication and pairing of double stranded DNA-helices, it is an exciting time for those interested in the biochemistry of specialized DNA structures<sup>[24]</sup>.

### 2.1.2. *Natural Hoogsteen-bonded triplexes.*

Naturally occurring homopurine/homopyrimidine sequences can fold into triplex configuration by binding a third strand of DNA or RNA in the major groove of W-C duplex DNA, through Hoogsteen or reverse Hoogsteen hydrogen bonds<sup>[25]</sup>. The purine strand of the W-C duplex engages the third strand through Hoogsteen hydrogen bonds in the major groove while maintaining the original duplex structure in a B-DNA like conformation. Triplexes offer exciting perspectives both toward oligonucleotide-directed control of gene expression and their formation in vivo as H-DNA. The latter is believed to be involved in diverse DNA transaction processes including DNA replication, gene transcription, recombination, and genome instability<sup>[26]</sup>. In H-DNA a long homopurine/homopyrimidine sequence donates the third strand of DNA to an adjacent compatible double helical acceptor. Interestingly, in silico analysis suggests that sequence capable to adopting DNA structures are very abundant in the human genome<sup>[27]</sup>. In contrast to other DNA secondary structure-forming sequences, which are typically located in the intergenic or intronic regions, H-DNA forming sequences are found embedded most frequently in promoters and exons and have been shown to be involved in regulating the expression of several disease-linked as well as normal genes. If H-DNA structure were to occur in vivo, it is likely that their formation or dissociation would involve cellular proteins that specifically recognize triplex DNA. Recently, studies with anti-triplex DNA monoclonal antibodies have found evidence

for triplexes in metaphases chromosomes<sup>[27]</sup> and an inhibitory effect on cell proliferation when the antibodies are introduced into cells before mitosis. This lead to the hypothesis that triplexes and their binding proteins serve as a bridge between the mitotic fuse and chromosomal DNA to facilitate proper chromosome segregation.

## **2.2. pH and cation effects on the properties of DNA triplexes.**

Potential gene-targeted therapeutic and biotechnological applications have provoked considerable interest in triplex formation and stability. The formation pathway of a triplex is significantly pH dependent and fell in three groups: under acidic conditions, the triplex is formed by a one-step docking, under near physiological conditions, the Watson-Crick duplex is first structured and then accepts the Hoogsteen third strand into its major groove, and under basic conditions, the triplex is not formed. In general, parallel triplexes are expected to be more stable than anti-parallel triplexes. However, their pH dependence might limit their physiological stability. The pH dependent thermodynamics can be study via circular dicroism (CD) and UV analysis.

### **2.2.1. Protonation effect on the triplex thermodynamic**

To evaluate the possible effect of protonation on the different base pairs and backbones, the pH dependent thermodynamics of the Py\*Pu·Py triplex, the Py·Pu antiparallel duplex and the parallel Py\*Pu duplex were studied under identical conditions. These data indicate that the thermodynamic stabilities of the Py\*Pu·Py triplex and the Py\*Pu duplex are clearly pH-dependent and are enhanced with decreasing pH, while Py·Pu duplex shows a slight reduction under the same conditions. This suggests that protonation has two opposite effects: one is to promote the stabilities of the Py\*Pu·Py triplex and Py\*Pu duplex, the other is to decrease the stability of the Py·Pu duplex. In addition, under acidic conditions free energy of the global triplex is much smaller than the sum of the free energies of the free Py·Pu duplex and free Py\*Pu duplex. This suggests that a larger energy would be required to disrupt the H\*W·C triplex compared to individually disrupting the Py·Pu duplex and Py\*Pu duplex. Alternatively, the triplex is a more stable complex than expected and its unfolding would be the one-step collapse of the global molecule under acidic conditions.

### **2.2.2. Ionic strenght effects on triplex stability**

The stability of a triplex can be influenced by the ionic strength and mixture of monovalent and divalent cations. The effect of the Na<sup>+</sup> have been investigated<sup>[28]</sup>. Particularly, at low Na<sup>+</sup> and strand concentrations, the melting temperatures of the Hoogsteen transition are too small to be determined, and they are not osservable increasing only Na<sup>+</sup> concentrations. Although the melting temperature of both the Hoogsteen and Watson-Crick transitions increase with increasing Na<sup>+</sup> concentrations, suggesting that the stability of the Hoogsteen paired strand more strongly depends on the ionic strength than does that of the Watson-Crick paired duplex. Moreover, if the third strand dC residues in the triplex were uncharged at pH=7, then one would expect a positive ionic strength due to the high negative charge density arising from the negatively charged phosphates on the three strands. Such ionic strength dependence of stability is the norm for duplexes and triplexes with uncharged bases, e.g. poly U\*A·U. In contrast, inverse salt dependence of T<sub>m</sub> at moderate ionic strengths is observed with duplexes stabilized by ionic interactions, as in the case of poli C<sup>+</sup>·C and poli- A<sup>+</sup>·A<sup>+</sup>. Similarly, when the NaCl concentration increase above 0.9M, third strand binding results destabilized. This reversal of salt dependence to increased shielding by the Na<sup>+</sup>, resulting in a reduced affinity of the dC residues for H<sup>+</sup> and



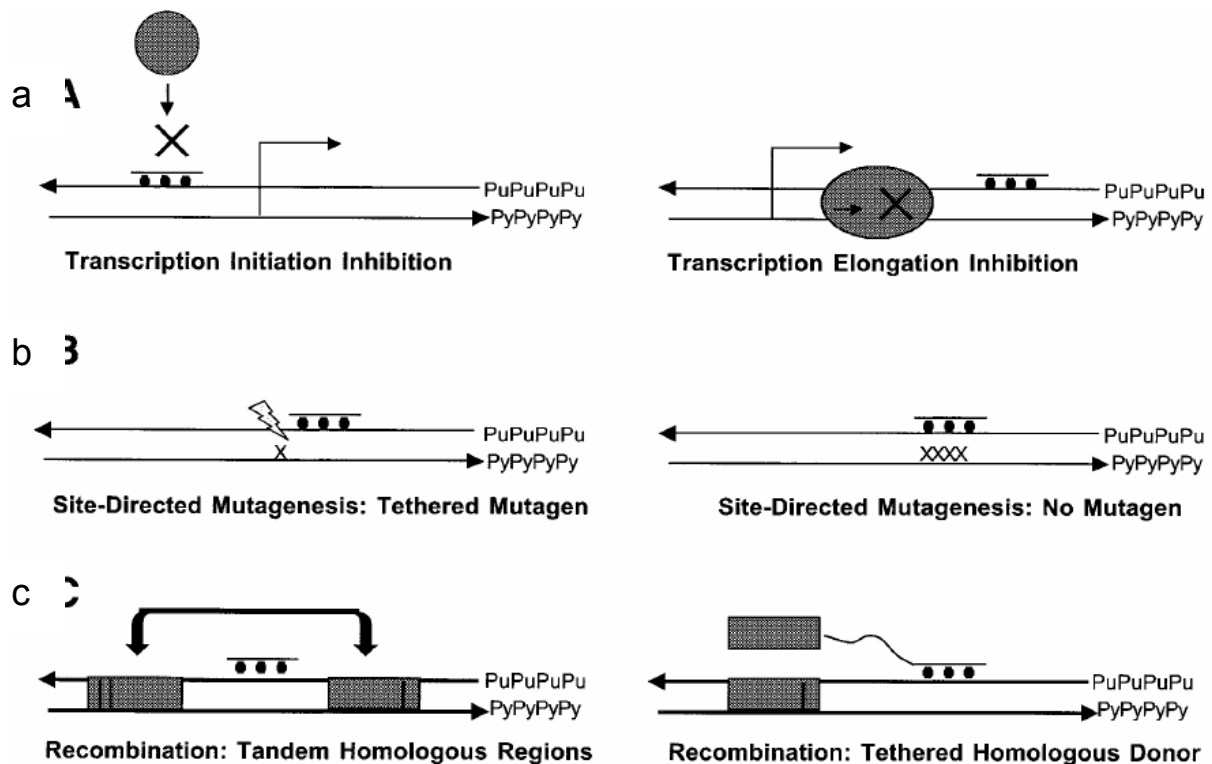
a decrease in the electrostatic attraction between the C<sup>+</sup>-G Hoogsteen base pair and also the dC<sup>+</sup> residues and the phosphodiester backbone of the duplex.

### 2.2.3 *Cation effect on the triplex thermodynamic*

The stability of the negatively charged DNA is explicitly dependent on the counterions. The influence of various cations, such as an alkali metal (Na<sup>+</sup>), alkali-earth metals (Mg<sup>2+</sup>, Ca<sup>2+</sup> and Ba<sup>2+</sup>) and transition metals (Mn<sup>2+</sup>, Co<sup>2+</sup>, and Zn<sup>2+</sup>), have been investigated<sup>[29]</sup>. The data indicate that the thermodynamic stability of a triplex is much greater with divalent cations than with monovalent cations, and almost a 100-fold concentration of Na<sup>+</sup> is required to stabilize the triplex as effectively, as Mg<sup>2+</sup> or Mn<sup>2+</sup><sup>[30]</sup>. This may be related to the valence-specific cationic stability of the triple helix structure. In the series of Mg<sup>2+</sup>, Mn<sup>2+</sup>, Ca<sup>2+</sup>, and Ba<sup>2+</sup>, the Watson-Crick duplex stability in the triplex increases from left to right (Mg<sup>2+</sup> > Mn<sup>2+</sup> > Ca<sup>2+</sup> > Ba<sup>2+</sup>). This order is in good agreement with the ionic radii of the divalent cations ( $r_{\text{Mg}^{2+}} = 0.66 \text{ \AA}$ ,  $r_{\text{Mn}^{2+}} = 0.80 \text{ \AA}$ ,  $r_{\text{Ca}^{2+}} = 0.99 \text{ \AA}$ , and  $r_{\text{Ba}^{2+}} = 1.34 \text{ \AA}$ ) and suggests that the divalent cations with a smaller radius may increase the affinity of the nucleotide alignment, resulting in enhancing the stability of the double helix structure. Interestingly, the order of the divalent cations, for stabilizing the W-C duplex in the triplex, is consistent with that for stabilizing the free W-C duplex, but it is different from the Hoogsteen paired strand one. This suggests that the interaction of divalent cations stabilize the global triplex in an order nearly identical to that for stabilizing the host W-C duplex, due to the large difference in free energy between the W-C duplex and the Hoogsteen paired strand. The influence of divalent cations on Py\*PuPy triplexes shows a slight difference on the Pu\*PuPy triplexes. Since the protonation at N3 of Hoogsteen cytosine residues plays a key role in stabilizing the Py\*PuPy triplex structure, while the Pu\*PuPy triplex doesn't require this precondition, the interaction of divalent cations with a triplex may be very different for the two classes of triplexes. It has been reported that Mg<sup>2+</sup> is crucial for the formation of the Pu\*PuPy triplexes, but not for the formation of the Py\*PuPy triplexes<sup>[31]</sup>. Another possibility may be due to the differences in buffer species, pH value, and base sequences.

## 2.3. *Triplex applications*

The sequence specificity of triplex forming molecules and their ability to compete successfully with other DNA binders gives these molecules great potential as tools to alter gene expression. Permanent heritable genetic changes can be induced by recombination and mutagenesis, while gene expression can be regulated by inhibition of RNA transcription. In figure 2.3 are reported six key means by which TFOs can be used to directly alter gene expression. In figure 2.3a is illustrated how transcriptional processes can be stopped by two different strategies. Both strategies are based on the binding of TFO to a target site and the subsequent creation of a physical block to a normal cellular process. Figure 2.3b shows both the use of TFOs to deliver a tethered mutagen as well as the ability of a TFO to cause mutagenesis within its binding site, even when not conjugated to a DNA reactive molecule. Figure 2.3c highlights two examples of experimental protocols in which TFOs have been used to promote homologous recombination between repeat sequences, which can be induced by TFO binding alone, and homologous recombination of a target genome site with a donor fragment, that is localized to its target by the TFO.



**Figure 2.3:** TFO alteration of gene expression. (a) Blocking of transcription with transcription factors or by interference with elongation. (b) TFO directed site – specific mutagenesis. (c) Gene correction via homologous recombination.

### 2.3.1. Transcription

Altered gene expression can be made via inhibition of transcription of extrachromosomal reporter genes and it seems best achieved through the binding of regulatory rather than coding sequence. TFO can bind in regulatory regions and compete with the binding of transcription factors, thereby preventing the initiation of transcription. Studies have shown that TFOs can bind to their genomic target *in vitro*, and in cells within the chromatin structures<sup>[32]</sup>. *In vitro* studies have shown highly specific TFO binding at regulatory sites in physiologic conditions, in this way TFOs block transcription and/or inhibition of regulatory protein factors binding. Some studies show that triplex binding sequences, which are tightly associated with nucleosomes, cannot be targeted unless the nucleosome association is removed. There have also been some successes in targeting the coding region rather than the regulatory region of a gene, such as the block of various polymerases.

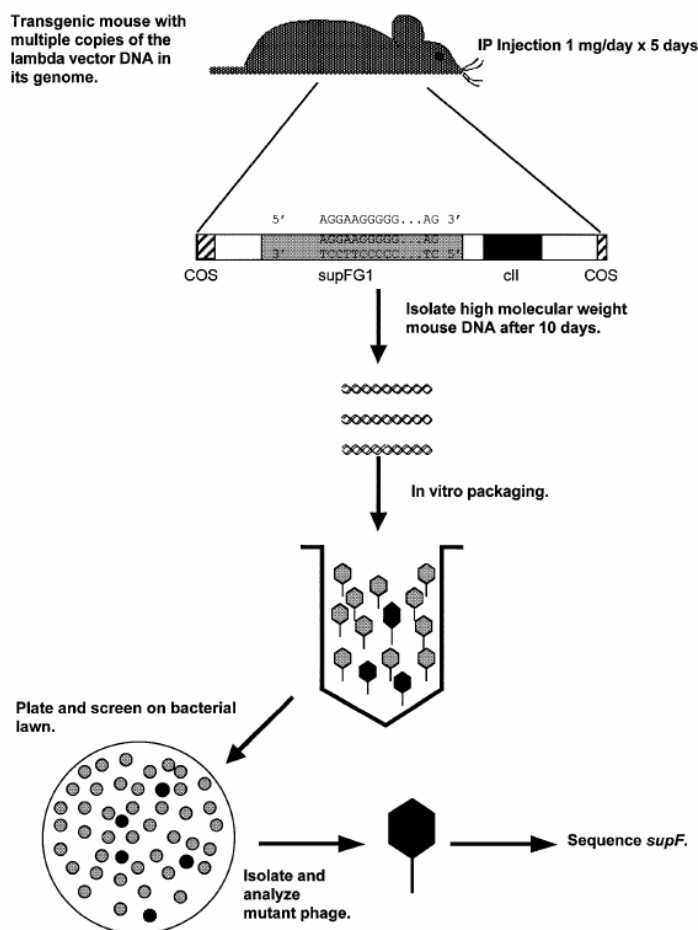
### 2.3.2. Genome modification

TFOs can also create permanent heritable changes in the genome. A number of studies have shown that TFOs are able to increase rates of site-specific mutagenesis<sup>[33]</sup> and site specific recombination<sup>[34]</sup>.

### 2.3.3. Triplex-directed mutagenesis

Both purine and appropriately modified pyrimidine TFOs have been used to direct mutagens to specific DNA sites. One such agent was psoralene. Psoralene-linked TFOs intercalate and bind covalently to duplex DNA when activated with long wavelength UVA irradiation increasing mutagenesis both *in vitro* ( plasmid targets) and in mammalian cell culture lines. Wang et al. observed that plasmide mutagenesis rates in cells were highly dependent upon binding affinity and delivery concentration. TFOs with Kds in the range of

$10^{-9}$  M were able to induce mutagenesis, whereas those with Kds of  $10^{-6}$  M were not. Studies with psoralene-linked TFOs also produced the first evidence that chromosomal sites could be targets for mutagenesis. More surprisingly is the discovery that in absence of psoralene, however, only purine TFOs result able to induce genomic changes (either mutagenesis or recombination) via the effects of third-strand binding alone. Recently Vasquez et al.<sup>[33]</sup> demonstrated that systematically administered TFOs could induce mutation at specific genomic sites in the somatic cells of adult mice (figure 2.4).



**Figure 2.4:** Experimental protocol to detect site-specific genome modification by systematically administered TFOs in mice.

In this work, a 30mer purine modified TFO containing a 3'-propanolamine group to prevent exonuclease-mediated degradation was injected intraperitoneally for 5 consecutive days. The TFO was targeted to chromosomal copies of a  $\lambda$ supFG1 reporter vector. After an additional 10 days, the mice were killed and tissues analyzed. The mice treated with the sequence specific TFO showed a 5-fold elevated mutation rate in the target gene  $\lambda$ supFG1, than in the non-targeted *cII* gene. All tissues tested resulted have mutation, except the brain. This is consistent with TFOs being unable to cross the blood-brain barrier.

#### 2.3.4. Triplex-induced recombination

Non-covalent triplexes formed by purine TFOs result active in stimulating repair and recombination in cell extracts, in cell culture and even in mice. Studies in human cell extracts and in DNA repair-deficient human cell lines have revealed the ability of non-covalent purine motif triplexes to provoke recombination in a manner dependent on the

nucleotide excision repair (NER) pathway<sup>[35]</sup>. In chromosome targeting work, two mutant thymidine kinase (TK) genes were integrated into a single chromosome site in mouse fibroblasts, followed by transfection of the cells via cationic lipids. High-affinity TFOs targeted against a region between the two TK-genes, stimulated recombination at a frequency of 6-7-fold above background and when the TFOs were microinjected into the nuclei of the cells, the yield increased up to 3000-fold<sup>[36]</sup>. These results suggested that TFOs could induce recombination at a chromosomal locus in mammalian cells, and highlighted the need for effective intranuclear delivery, to increase levels of targeted genome modification. The observation that the third strand binding can provoke DNA repair and stimulate recombination provides a strategy to mediate targeted gene conversion using a TFO linked to a short donor DNA fragment, figure 2.3c. The donor fragment was homologous to the target site, except for the base pair to be corrected<sup>[37]</sup>. In this bi-functional molecule, the TFO domain mediates site-specific binding which direct the donor to the desired gene. This binding is proposed to trigger, repair or sensitize the target site to recombination. Gamper et al.<sup>[38]</sup>, using a bi-functional oligonucleotide, consisting of a guiding TFO domain and a strand invading homology domain, demonstrated strand invasion of a supercoiled plasmid DNA in an in vitro study. Interestingly, physical TFO-linkage to the donor was unnecessary. A co-mixture of separate TFO and donor segments also yielded elevated gene correction frequencies. This latter result confirms TFO ability to promote recombination between the target site and a third, unlinked DNA molecule, if it is present in high concentration. However, mechanistic question remain unanswered. Though pathways have been implicated in the mutagenesis and recombination effects of TFOs, the exact mechanism are unknown. Studies occur to establish the precise cellular processes that play a role in mediating these genetic changes. It is also of interest to determine which factors can influence the availability of genomic targets to DNA binding molecules: transcription levels, cell cycle phase, nucleosome binding and histone state may all play important roles. A major long term goal of the field is to develop TFOs as tools for the target modification. Efficiency, delivery and genome availability must be reconsidered. The rate of mutagenesis remains at a fraction of a percent. Recently, techniques for the intracellular production of a single-stranded DNA molecule have been developed such as the use of vectors to produce TFOs intracellularly. However, disease, which require the majority of cells in a tissue to be corrected, may not be amenable to triplex technology. Instead, triplex technology is probably applicable to certain hematologic, metabolic and hepatic disease that requires a small percentage of the target cells to ameliorate the phenotype of the disease.

#### 2.3.5. Other applications: properties of functional molecules conjugates-TFOs

Oligonucleotides covalently linked with functional molecules have been attracting current interest because of their widespread biological usage. Interesting results have been obtained in the application of functional molecules conjugates-TFOs, that provide new potential tools for altering gene function. For example, Helene et al.<sup>[39]</sup> have reported that oligodeoxynucleotide (ON) attached to a phenanthroline-copper chelate at the 5'- termini cleaves DNA by triple-stranded formation showing high specificity. Miller et al.<sup>[40]</sup> have shown that methylphosphonate-ONs covalently linked to psoralen and targeted single-stranded DNA give cross-linked duplexes by cycloaddition reaction under photoirradiation. TFOs can be conjugated to non-specific antitumour drugs, such as camptothecin<sup>[41]</sup> that normally induce massive non-specific DNA damage in cells, converting topoisomerase I into a cell poison by blocking the relegation step, thereby enhancing the formation of persistent DNA breaks responsible for cell death. Triplex-forming oligonucleotide-camptothecin conjugates (TFO-CPT) can bind in a sequence specific manner to the major

groove of double helical DNA and positions the drug selectively at the triplex site, thereby stimulating topoisomerase I-mediated DNA-cleavage at this site. TFOs are also designed to contain a chelating agent, such as a bipyridine moiety<sup>[42]</sup>. Transition metal complexes-containing oligonucleotides can be involved in the study of electron transfer processes, as well as in the development of artificial nucleases characterized by a high sequence-specificity and efficiency. In fact, a metal centre tethered to a particular sequence could enable the complex to oxidatively modify or cleave a nucleic acid target. However, drug-conjugated TFOs can be used to enhance drug-selectivity and to improve lower toxicity. Daunomycin-conjugated TFOs (dauno-TFOs) are able to down regulate promoter activity and transcription of the target endogenous gene. Particularly, Myc-targeted-dauno-TFOs inhibits growth and induced apoptosis of prostate cancer cell constitutively expressing the gene, in a sequence specific and triplex mediated manner. To test the selectivity of dauno-TFOs, they have been examined on growth of normal human fibroblasts, which express low levels of c-myc. Despite their ability to inhibit c-myc transcription, dauno-TFOs failed to inhibit growth of normal fibroblasts at concentration that inhibited growth of prostate cancer cells. In contrast, daunomycin inhibited equally fibroblasts and prostate cancer cells<sup>[43]</sup>. These data indicate that dauno-TFOs are attractive gene-targeting agents for development of new cancer therapeutics.

## **2.4. Recent improvements in antigene technology**

Selective artificial control of gene expression is a long-standing dream in biotechnology and human therapy. Oligonucleotides seem perfectly suited for this purpose because of their unique base-base recognition properties. The potential advantages in targeting genomic DNA by oligonucleotides are counterbalanced by several restrictions common to all oligonucleotide-based approaches (e.g. bioavailability, cell permeability and non-specific binding of the oligonucleotides or analogues used). Further, the triplex approach shows specific restrictions. Some of them are motif-dependent while others are of general nature. For the parallel motif, the severest restriction is the pH-sensitivity of the C<sup>+</sup>\*G<sup>-</sup>C base triplet, which is intrinsically unstable under physiological conditions due to the low pKa of the third strand cytosines. Triplex stability in the antiparallel motif is markedly sequence-dependent due to alternative structure formation (e.g. G tetrads), of the corresponding GA or GT rich TFOs under physiological conditions (especially K<sup>+</sup> concentration). Finally there are common problems in both motifs. The most prominent ones are related to target accessibility, low overall affinity of TFOs to their target, resulting in poor competition with DNA binding proteins, and the restriction of the known binding motifs to homopurine/homopyrimidine DNA sequence tracts.

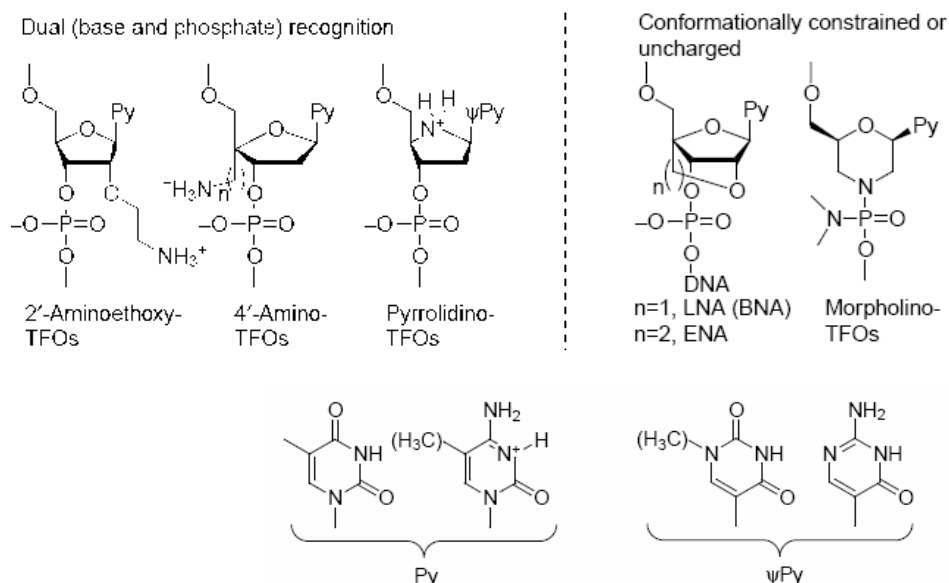
### **2.4.1. Increasing triplex stability**

There are a few basic concepts in improving the binding of TFOs to duplex targets that have been successfully applied in the past few years to increase triplex stability. First, the introduction of positively charged amino functions on the sugar and/or base unit, designed to make additional binding contacts between phosphodiester residues of the target Watson-Crick duplex besides the base-base interactions. Second, the use of conformationally adapted or constrained TFOs. Finally, the use of non-charged TFOs. Most recent progress has focused on the parallel binding motif of TFOs.

### **2.4.2. Dual recognition**

The best explored and most promising parallel motif TFOs are those containing an aminoethoxy side-chain at C(2') of the ribonucleosides units, developed by Cuenoud and

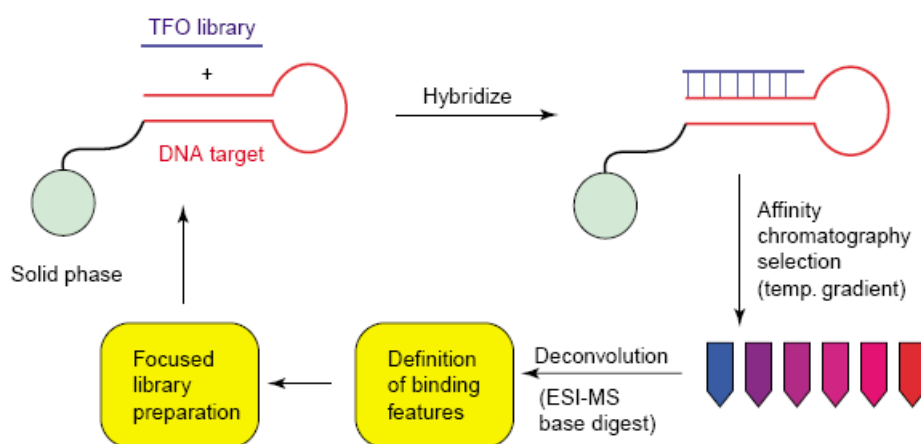
colleagues<sup>[44]</sup>. This amino function results protonated under physiological pH, leading to an increase of c.a. 3.5°C per modification at pH=7.0 Fully modified 2'-aminoethoxy TFOs (AE-TFOs, figure 2.5) can lead to triplexes more stable than the target duplex alone. This kind of modification also improve a gain in thermodynamic stability ( $\Delta G$ ) of c.a 0,5 Kcal/mol for each modified unit that results in a 1000-fold enhanced association rate compared with that of a DNA-TFO. This is an important feature, considering the lower association kinetics for triplex formation than duplex formation. Further, the 2'-amino side-chain occurs in a gauche<sup>+</sup> conformation and shows astonishingly low conformational mobility in the triplex<sup>[45]</sup>. Studies on variations in length of the side chain indicate they do not improve binding efficiency.



**Figure 2.5:** Molecular structures of TFOs with potential dual recognition properties, with conformationally constrained sugar analogues and with uncharged backbone structures.

Another point of modification using alkylamino chains is the C(4') of the deoxyribose unit. 4'-amino-TFOs have been synthesized by Matsuda and co-workers<sup>[46]</sup>. They found the same thermal triplex stabilization than for the ethylamino side chain, by than 1°C per modification. Shortening or lengthening the tether between C(4') and the amino function consistently reduced affinity relative to the two-carbon linker. However, the efficiency of this system is not better than that of AE-TFOs. Within this context, the pyrrolidino-pseudo-nucleosides have been proposed. According to modelling based on X-ray data available for parallel motif triplexes, the furanose ring oxygen is close to a pro-R non-bridging phosphate oxygen of the purine duplex strand. Experiments with TFOs containing uracil or N-1-methyl-uracil for recognizing adenine in the parallel binding motif, however, showed triplex destabilization, whereas those containing the base isocytosine increased triplex stability at neutral pH by c.a. 2°C in melting temperature ( $T_m$ ) per modification. The reason for the different stabilization/destabilization improved by the pyrrolidino-pseudo-nucleosides is unclear. Another interesting consideration is that RNA-TFOs generally form more stable triplexes than do DNA-TFOs. NMR analysis shows that deoxyribose units in DNA-TFOs don't adopt the same sugar pucker<sup>[47]</sup>. Thus, the overall helical conformation of DNA triplexes is somewhere between A and B. A recent study with 2'-deoxy- and 2'-O-methyl-RNA-TFOs showed marked sequence-dependent differences in stability, that cannot easily be rationalized with known rules in conformational analysis<sup>[47]</sup>. However, a combinatorial method to determine conformational effects of TFOs on triplex stability was developed<sup>[48]</sup>. A

schematic representation of this approach is reported in figure 2.6. Within a given base recognition frame imposed by the target duplex, an oligonucleotide library<sup>[49]</sup>, containing either a ribo- rU/rC or a deoxyribo-dT/dC sugar unit at each position in the chain, was constructed by the split and combine technique. This library was then loaded on the agarose-supported DNA target and fractionated via affinity chromatography in a temperature gradient. The fractions were then deconvoluted by mass-spectrometry, combined with alkaline digestion, to determine the preferred composition and location of ribonucleosides within the sequences. This information needs to produce focused libraries.



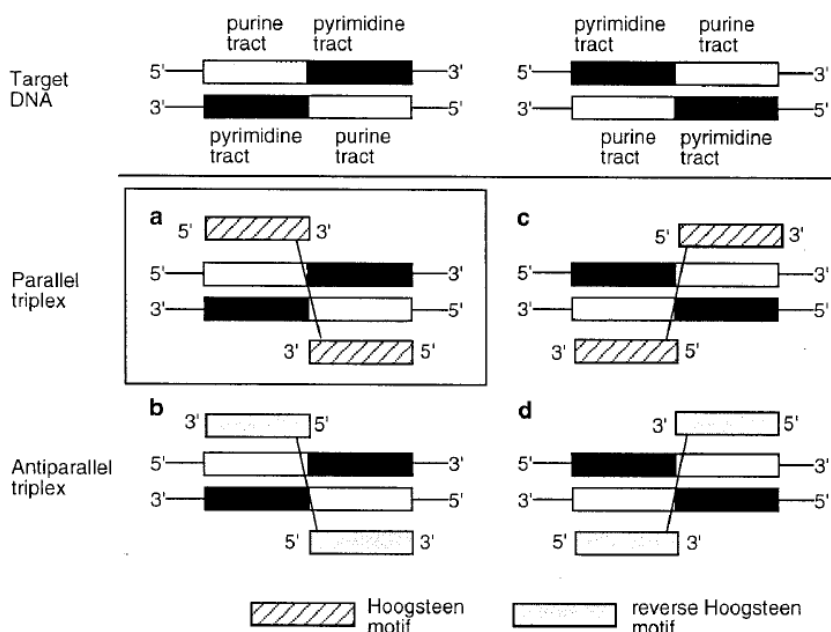
**Figure 2.6:** Schematic representation of the combinatorial approach developed for establishing a structure/stability profile of TFOs with conformational variation in the backbone.

Another interesting approach involves the use of TFOs containing conformationally constrained nucleoside units. Given the preference of RNA- over DNA-TFOs in the parallel binding motif, oligonucleotide analogs that are constrained in a 3'-endo conformation (RNA like), should be interesting. Recently a triplex containing oligopyrimidine nucleotides partially modified with 2'-O,4'-C-methylene nucleosides (figure 2.5), which are known as bridged nucleic acids (2',4'-BNA) or locked nucleic acids (LNA)<sup>[50]</sup> was greater than that of a triplex containing natural 2'-deoxyribose units at physiological pH. An increase in the association constant, which is due to a decrease in the dissociation rate, was observed. Surprisingly, fully modified LNA-TFOs don't form triplexes with DNA targets, there exist an optimal number of LNA-residues within a DNA-TFO that exert a stabilizing effect. Data also exist on triplex binding with LNA-analogue ENA<sup>[51]</sup>, containing a 2',4'- ethylene bridge. ENA is able to stabilize triplex formation at neutral pH roughly to the same extent as LNA and it is more nuclease-resistant than natural DNA and LNA. Furthermore, fully modified ENA-TFOs bind to their DNA target, differently from fully modified LNA-TFOs that do not. The reason of this phenomenon is not completely understood.

#### 2.4.3. Alternate-Strand Triple helix formation

In the antigene strategy, proposed to control gene transcription, an oligopyrimidine-oligopurine sequence of double helical DNA is recognized by a third-strand oligodeoxynucleotide that binds to the major groove of the DNA and forms a local triple helix (triplex). Target sequence in the antigene strategy are very restricted. Since the thermal stability of the triplexes is generally lower than that of the duplexes under physiological conditions, an oligopurine cluster with fairly long chain lengths is required for stable triplex formation. In addition, formation of the pyrimidine motif triplex needs conditions of low pH (<6), because unmodified cytosines residues, if present in the third strand, must be protonated to bind with guanine of the G-C duplex. Several approaches

have been described to overcome this limitation. If the target duplex is composed of two adjacent and alternating oligopurine-oligopyrimidine tracts, the duplex can form a triplex with a single strand ON which pairs with oligopurine tracts, recognition can be achieved by a single third strand, portions of which are targeted against single oligopurine tracts. Such an oligonucleotide zigzags along the major groove, switching from one oligopurine strand to another at the 5'-purine-pyrimidine-3', or 5'-pyrimidine-purine-3'-step<sup>[52]</sup>. This so called "alternate-strand-triplex" or "switched triplex" can be formed in two distinct ways. By combining both Hoogsteen and reverse Hoogsteen motifs, the natural ONs with oligopurine-oligopyrimidine sequences such as 5'-oligopurine-oligopyrimidine-3' or 5'-oligopyrimidine-oligopurine-3' can be used as the third strand. In the second approach, only one set of motifs (Hoogsteen or reverse Hoogsteen motif) is employed, namely, the 3'- and 3'- or 5'- and 5'- ends of two ODN fragments have to be connected using a suitable linker. In this latter case the different possibilities for "alternate-strand-triplex" formation are shown in figure 2.7. When the oligopyrimidine sequence follows the oligopurine sequence, the 3'-3'-linked and the 5'-5'-linked ONs can form the parallel (a) and antiparallel (b) triplexes with the target duplex by Hoogsteen and reverse-Hoogsteen hydrogen bonds, respectively. On the other hand, when the oligopurine sequence follows the oligopyrimidine sequence, the 5'-5'-linked and the 3'-3'-linked ONs can form the parallel (c) and antiparallel (d) triplexes with the target duplex by Hoogsteen and reverse-Hoogsteen hydrogen bonds, respectively. Several 3'-3' or 5'-5' internucleoside junctions have been proposed as linkers for alternate strand TFOs, with the latter imparting a minor cooperativity in the binding of the two domains of the TFO with the target duplex. From a chemical point of view, the 3'-3' or 5'-5' inversion can be fulfilled by a suitable linker capable of crossing the major groove and whose properties can be, in addition, exploited to supply the oligonucleotide with useful characteristics.

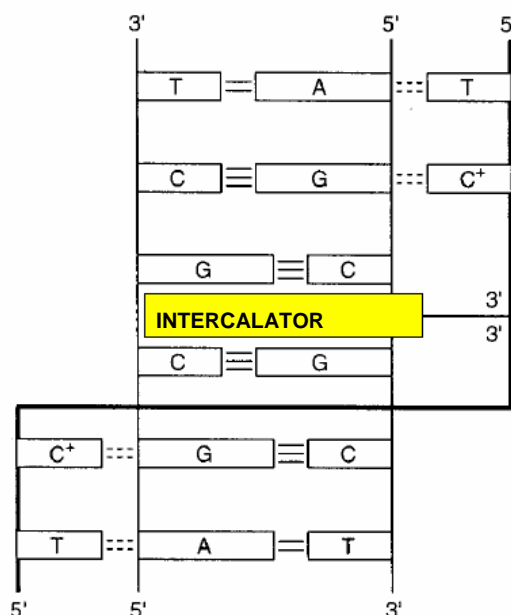


**Figure 2.7:** Schematic presentation of "alternate-strand triplex" formation by the 3'-3'- or 5'-5'-linked ONs.

For example, the 3'-3' linker may incorporate an intercalating agent or a major groove ligand in order to improve the hybridization between the probe ON and the target duplex and to enhance the thermal stability of the triplexes. In most recent years the synthesis of ONs containing a 3'-3' inversion polarity motif (figure 2.8), able to hybridize the target



duplex, in parallel mode, by formation of Hoogsteen triplets<sup>[52-54]</sup> gained much attention. Results on the stability of these triplex structures pointed to the phosphodiester bond as a suitable 3'-3' junction into TFOs.

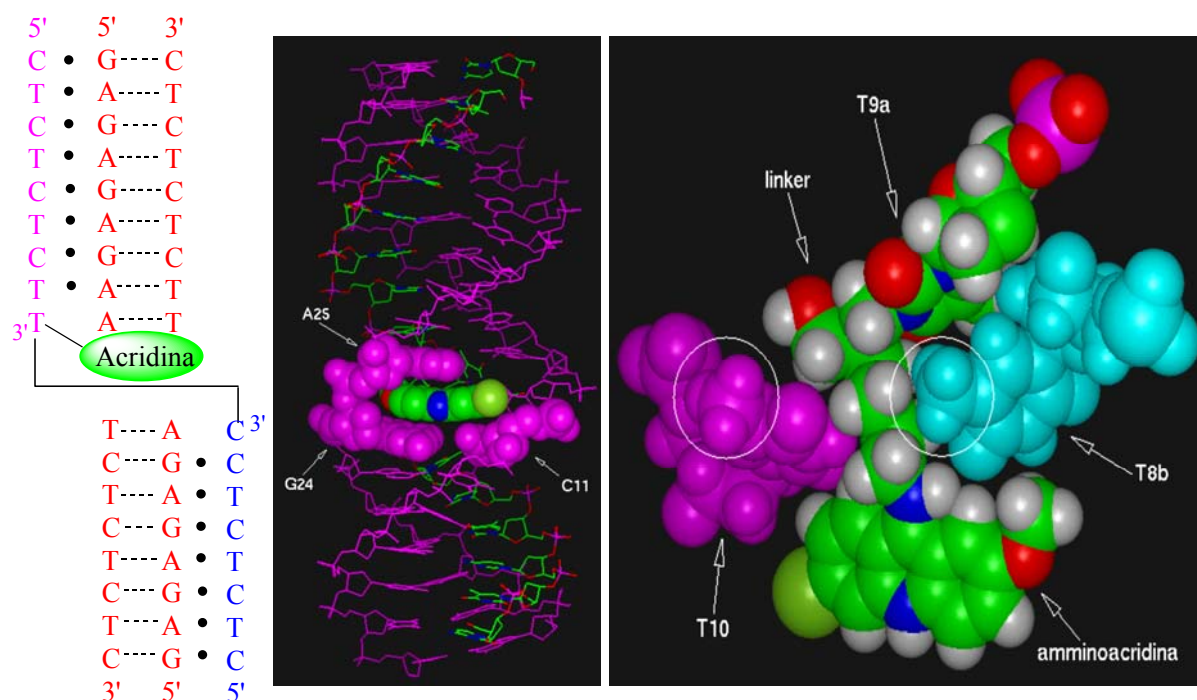


**Figure 2.8:** 3'-3'-linked ON with an intercalater at the junction point.

Nevertheless, molecular mechanic calculation indicated that a distortion of the sugar-phosphate backbone around the 3'-3' phosphodiester junction occurs thus preventing the correct Hoogsteen hydrogen bonds formation for the nucleotide bases flanking the site of inversion polarity<sup>[53]</sup>. In order to balance destabilizing effects occurring in alternate strand triplexes, an intercalating agent near the 3'-3' switching region of the TFO<sup>[52-54]</sup> was introduced.

## 2.5. The aim of the work

The aim of this work is the conjugation of a polypyrimidinic TFO with an acridine residue in the proximity of the 3'-3' inversion site, so that its intercalation may counterbalance possible destabilizing effects of the junction (figure 2.9). Among the possible DNA intercalator agents, acridine is particularly advantageous considering that: i) molecules containing an acridine moiety have been extensively studied as duplex and triplex DNA stabilizing agents; ii) acridine linked ONs have been proposed as antisense agents and as efficient TFOs;<sup>[52-54]</sup> iii) ONs linking acridine derivatives can induce a DNA cleavage under photoirradiation when combined with metal ions;<sup>[53-54]</sup> iv) a large number of acridine derivatives are commercially available. It is also to be noted that the only synthetic strategy so far proposed to prepare such ON analogues is very complex. This work would provide an easy and convenient method for the synthesis of TFOs carrying a chloro-metoxiacridine residue attached to the N-3 of a thymidine flanking the 3'-3' phosphodiester bond, through an alkylamino linker.



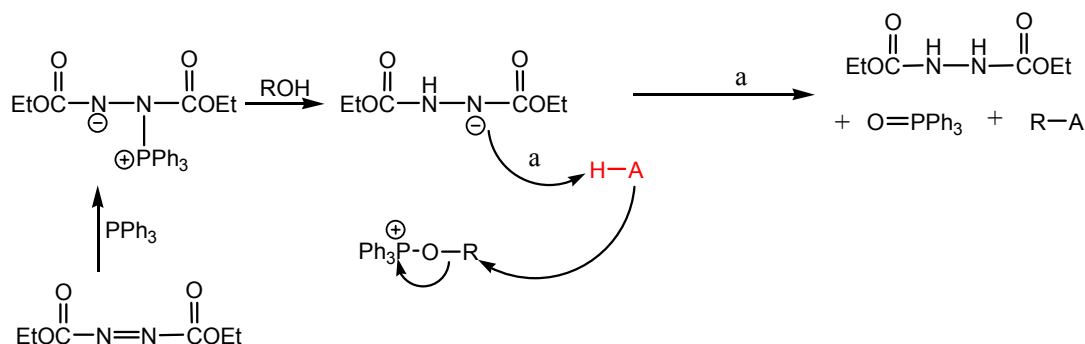
**Figure 2.9:** 3'-3'-linked ON 5'-(CT)<sub>4</sub>T<sup>acridine</sup>-3'-3'-C(CT)<sub>3</sub>C. Molecular modelling.

To realize it, the N-3 position reactivity of thymidine towards alkylation was considered, together with the reported stability of duplex and triplex DNA complexes containing intercalators, or other groups, attached to N-3 position of a thymidine residue<sup>[53-54]</sup>. Using this simpler synthetic alternative, oligonucleotides with a 3'-3' inversion of polarity, containing an acridine moiety attached to the nucleotide base flanking the 3'-3' phosphodiester bond, have been prepared, characterized and used to study their hybridization properties in the formation of stable and non canonical triple helix structures. These complexes have been investigated by UV melting studies and CD experiments.

## 2.6. Chemistry

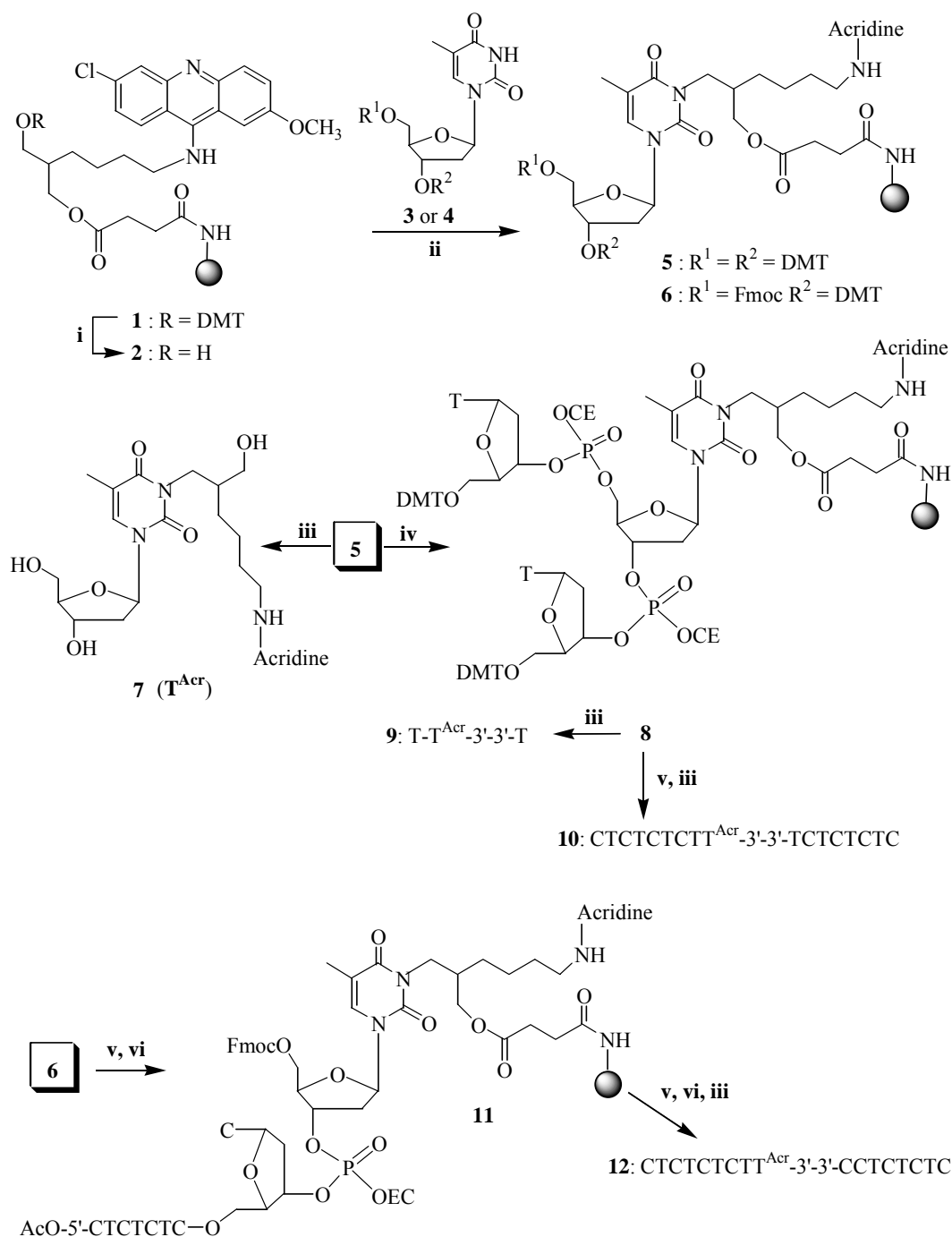
This approach is based on a commercially available difunctionalized CPG support carrying a 2-methoxy-6-chloroacridine moiety and a primary alcoholic hydroxy function protected with the DMT group (**1**, Scheme 2.1). A suitably 3',5'-protected thymidine **3** or **4** was linked to support **2**, obtained by removal of DMT protecting group, through the N-3 atom of the base using Mitsunobu condensing conditions<sup>[55-56]</sup> (figure 2.10), thus affording supports **5** and **6**, respectively (70% yield by DMT spectrophotometric test). Supports **5** and **6** were found to be stable under the chemical treatments required by the DNA chain assembly and cleavable, at the succinate function, in the final alkaline treatment for detachment and deprotection procedure. The cleavage from the support releases a thymidine base connected to the acridine residue through a hexylamino chain.

Support **5**, after capping of the unreacted hydroxyl functions and successive removal of 3',5'-DMT groups, reacted almost quantitatively with the 5'-DMT-thymidine-3'-phosphoramidite under standard automated procedure, giving support **8**.



**Figure 2.10:** Mitsunobu mechanism

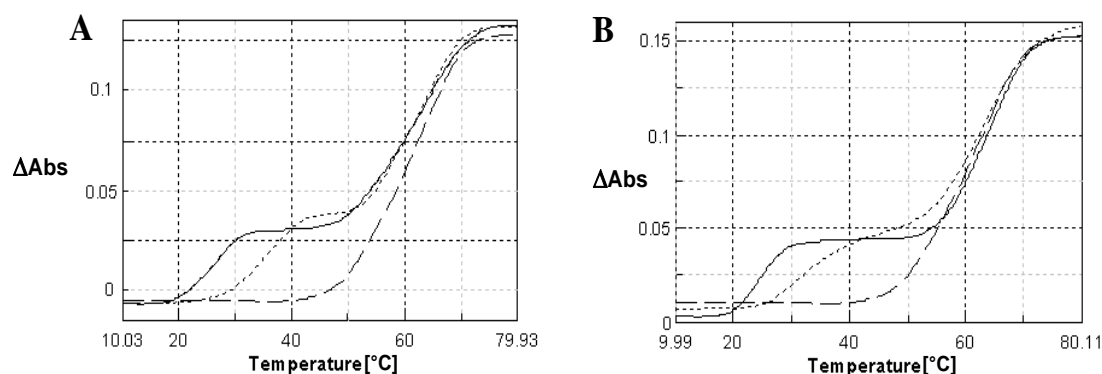
The deprotection step of supports **5** and **8**, by trityl removal, followed by an alkaline treatment with 0.4M solution of NaOH, furnished the expected acridine-nucleoside **7** and the acridine-trinucleotide **9**, respectively, the latter containing the 3'-3' phosphodiester junction. The structures of both released products were confirmed by  $^1\text{H}$  NMR and MS spectrometry. Eight coupling cycles performed on support **5**, with appropriate 3'-phosphoramidite nucleotide units, furnished TFO **10**. Support **6**, in which the 5' and 3'-OH functions are orthogonally protected with Fmoc and DMT groups, respectively, allowed the sequential synthesis of different ON domains of the TFO. After the removal of 3'-DMT group from support **6**, the elongation of the first ON domain by the formation of the 3'-3' phosphodiester junction, was achieved, thus leading to **11**. Then, the 5'-Fmoc group was removed by  $\text{Et}_3\text{N}$  /pyridine treatment and the second ON domain was assembled, thus furnishing the TFO **12**. Between the two alternative 5- or 3'-OH Fmoc protections, the former was chosen on the basis of preliminary experiments performed on solid supports carrying 5'- or 3'-Fmoc-thymidine. These had shown that complete 5'-deprotection could be achieved by 5% triethylamine in pyridine treatment in 1.5 hours whereas a prolonged reaction time was needed to quantitatively deprotect the 3'-hydroxy function. TFO **10** and **12** were synthesized on a 1  $\mu\text{M}$  scale by a DNA synthesizer. Crude oligomers, detached from the support as described for trinucleotide **9**, were purified by HPLC on an anionic exchange column. The collected peaks were desalted thus furnishing more than 95% pure **10** and **12** (by RP-HPLC) in 28 and 24 units  $\text{OD}_{260}$  units yields, respectively. MALDI TOF-MS spectra of **10** and **12** were in agreement with their structures. Further, UV-VIS spectra of **10** and **12** confirmed the presence of the acridine moiety into the oligomers, showing a diagnostic absorption at  $\lambda_{\text{max}}$  424 nm.



**Scheme 2.1:** Reagents and conditions: i) TCA 3% in DCM; ii) TPP, DEAD, THF, DCM; iii) NaOH 0,1 M in (CH<sub>3</sub>OH-H<sub>2</sub>O 4:1 v/v); iv) coupling with thymidine; v) pyridine/Et<sub>3</sub>N (95:5 v/v, 1.5 h); vi) chain elongation with 3'-phosphoramidite deoxynucleoside building blocks.

## 2.7. UV and circular dichroism (CD) studies

The triplexes **14** and **15** (Table 2.1) were formed by mixing equimolecular amounts of TFO **10** and **12** respectively with the target duplex DNA **13**, and successively heating at 90° C for 5 min. The solutions were then equilibrated for 15 h at room temperature before performing the analyses. The formation and stability of the triplexes **14** and **15**, compared with the acridine devoid triplexes **16** and **17**, were studied by UV thermal denaturation experiments and by CD in 5 mM NaH<sub>2</sub>PO<sub>4</sub>, 140 mM KCl and 5 mM MgCl<sub>2</sub> buffer (at pH 5.5, 6.0 and 6.6). Figure 2.11 shows melting profiles of the duplex **13** and triplexes **14-17**. The T<sub>m</sub> values are reported in Table 2.1. Melting curves for **14-17** showed a typical biphasic behaviour, with the first sigmoid attributable to the triplex dissociation, while at higher temperatures the expected pH-insensitive duplex to single strands transition could be observed. The UV melting data indicated for the acridine containing triplexes **14** and **15** a stabilization effect at all the pH values with respect to the corresponding acridine devoid triplexes **16** and **17**. At pH 6.6 the presence of the acridine in the TFO led to a stabilization with ΔT<sub>m</sub> values of 8.5 and 7.9 °C, respectively, for triplexes **14** and **15**, while at pH 5.5 slightly lower effect were observed (ΔT<sub>m</sub> 6.8 and 6.6 °C, respectively).



**Figure 2.11:** UV melting profiles (pH 6.6) (A) triplex **14** (....), triplex **16** (—) and duplex **13** (- -) (B) triplex **15** (....), triplex **17** (—) and duplex **13** (- -)

Complexes	pH 5.5	pH 6.0	pH 6.6
Duplex <b>13</b>	62.3	63.2	63.4
Triplex <b>14</b>	41.4	36.7	34.5
Triplex <b>15</b>	39.4	34.9	32.5
Triplex <b>16</b>	34.6	29.1	26.0
Triplex <b>17</b>	32.8	27.0	24.6

5' -GAGAGAGAATCTCTCTC-3'  
3' -CTCTCTCTTAGAGAGAG-5'

### Duplex **13**

5' -CTCTCTCTX  
5' -GAGAGAGAA TCTCTCTC-3'  
3' -CTCTCTCTT AGAGAGAG-5'  
LYCTCTCTC-5'

**Triplex 14** : X = T<sup>Acr</sup>; Y = T

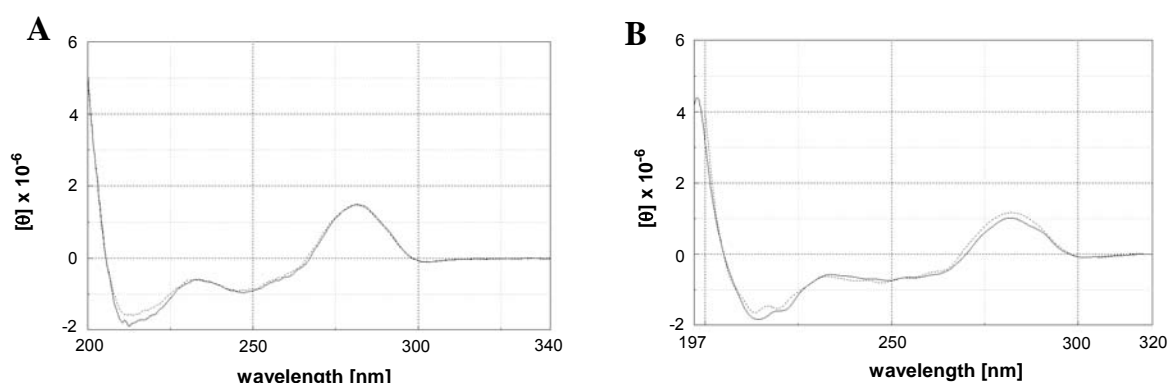
**Triplex 15** : X = T<sup>Acr</sup>; Y = C

**Triplex 16** : X = T; Y = T

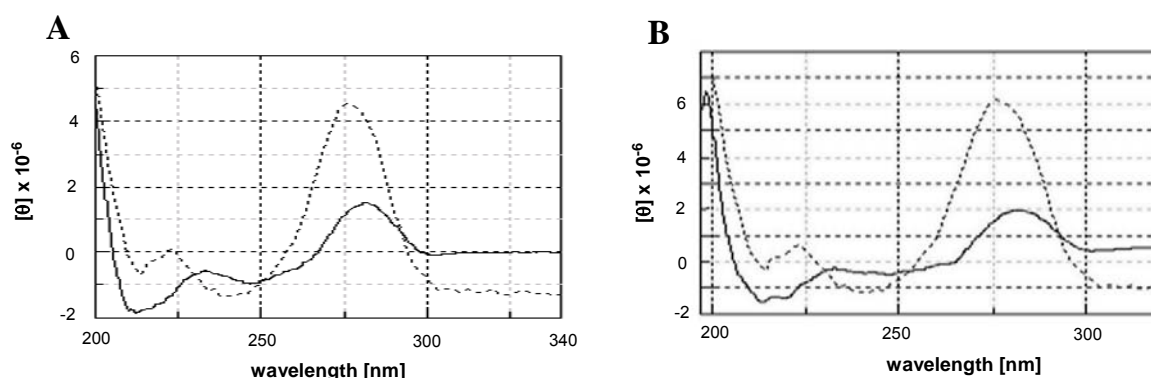
**Triplex 17** : X = T; Y = C

**Table 2.1:** T<sub>m</sub> values (C°)/pH for duplex **13** and triplexes **14-17**

The  $T_m$  values of the triplex **14** resulted to be higher than those of **15** at all the tested pH and the same behaviour was observed for acridine devoid triplexes **16** in comparison with **17**. These last data indicate that the substitution, into TFO, of a thymine (in **14** and **16**) with a cytosine (in **15** and **17**) at the 3'-3' junction, induces, as expected, a destabilizing effect, likely due to the electrostatic repulsion between consecutive protonated cytosine. The formation of a triplex structure for **14-17** was confirmed by circular dichroism measurements. The concentration of the triplex solution was in the range  $1.9\text{-}2.3 \times 10^{-5}$  M. CD spectra of triplexes **14-17** (pH 6.6) are reported in figure 2.12. Complexes **16** and **17**, as already reported for the same kind of complexes showed a CD profile very similar to those observed for canonical Hoogsteen type triplexes, characterized by a large positive band at 280 nm and a negative band centred around 213 nm. Particularly the negative band is indicative of the existence of a triplex structure. Almost the same patterns were observed for acridine containing triplexes **14** and **15**. The CD spectra of triplexes **14** and **15** were compared with the normalized summed spectra of the duplex **13** and appropriate single strand TFOs (Figure 2.13). Significantly, for **14** and **15**, the diagnostic band at 213 nm results deeper if compared to the corresponding band in the normalized sum spectra and the 282 nm band decreases in intensity and shows a red shift. These data are clearly indicative of a triplex structures for the complexes **14-17**.



**Figure 2.12:** (A) CD spectra at pH 6.6 of triplex **14** (—) and triplex **16** (...); (B) triplex **15** (—) and triplex **17** (...).



**Figure 2.13:** (A) CD spectra at pH 6.6 of triplex **14** (—) and normalized summed spectra (...) of duplex **13** and TFO **10**; (B) CD spectra at pH 6.6 of triplex **15** (—) and normalized summed spectra (...) of duplex **13** and TFO **12**.

## 2.8. Conclusions

This research work would provide an alternative easy synthesis of ONs containing a 3'-3'-phosphodiester linkage and bearing an acridine residue on the thymidine base flanking the 3'-3' junction. This synthesis was based on the preparation of a new kind of nucleoside-acridine solid support (**5** or **6**). This synthetic strategy could be an useful entry to link an intercalating agent near to 3'-3' (or 5'-5') phosphodiester linkage, thus furnishing a convenient method to stabilize alternate strand triple helices by minimizing the mismatching effect into the region between purine and pyrimidine domains. In fact, both CD and UV melting data, indicate that the acridine moiety, linked through a 7 atoms spacer arm to the N-3 of a thymine, does not hamper the formation of a triplex structure. Furthermore, the stabilization effect observed for triplex **14** and **15** strongly suggests an intercalation of the acridine residue into the triplex structure (figure 2.9).

## 2.9. Experimental Session

### 2.9.1. General Methods

The following abbreviations were used throughout the test: triplex forming oligonucleotide (TFO), fluorenylmethoxycarbonyl (Fmoc), 4,4'-dimethoxytrityl (DMT), trichloroacetic acid (TCA), dichloromethane (DCM), triethylammonium bicarbonate (TEAB), diethylazodicarboxylate (DEAD), triphenylphosphine (TPP), tetrahydrofuran (THF). Acridine CPG support was purchased from Glen Research. Functionalization of solid supports was carried out in a short column (5 cm length, 1 cm i.d.) equipped with a sintered glass filter, stopcock and a cap. The ON were assembled with a PerSeptive Biosystems Expedite DNA synthesizer using phosphoramidite chemistry. HPLC analyses and purification were carried out on a Waters 600E apparatus equipped with UV detector. UV spectra e thermal denaturation experiments were run with a Jasco V 530 spectrophotometer, equipped with a Jasco 505T temperature controller unit. CD spectra were obtained with a Jasco 715 circular dichroism spectrophotometer. NMR spectra were recorded with a Bruker AMX500 spectrometer. ESI mass spectrometric analyses were performed on a API 2000 (Applied Biosystem) machine used in negative mode. MALDI TOF mass spectrometric analysis was performed on a PerSeptive Biosystems voyager-De Pro MALDI mass spectrometer using a picolinic/3-hydroxy-picolinic acids mixtures as the matrix.

### 2.9.2. General Procedures

**5'-O-fluorenylmethoxycarbonyl-3'-O-(4,4-dimethoxytriphenylmethyl)-thymidine (4)** <sup>1</sup>H-NMR (acetone-d<sub>6</sub>) δ: 9.90 (1H, bs, N(3)-H); 8.02 (1H, s, H-6); 7.90-6.80 (21H, m's, aromatic protons); 6.32, (1H, m, H-1'); 4.52 (2H, m, CH<sub>2</sub> Fmoc residue); 4.32 (1H, m, H-3'); 4.24 (1H, t, CH Fmoc residue); 3.98 (3H, m, H<sub>2</sub>-5' and H-4'); 3.81 (6H, s, OCH<sub>3</sub> DMT); 1.92 and 1.72 (1H each, m's, H<sub>2</sub>-2'); 1.65 (3H, s, CH<sub>3</sub>-5);

#### Preparation of supports **5** and **6**; product **7**

200 mg of commercially available support **1** (0.044 meq/g) were washed with DCM and then treated with 3x 2.0 mL of a DCM solution of TCA (3%, w/v) for 2 min. This treatment was repeated three times and the resulting support **2** was exhaustively, washed with DCM and then dried under reduced pressure. 207 μL (1.3 mmol) of DEAD were added to 340 mg (1.3 mol) of triphenylphosphine (TPP) dissolved in 800 μL of THF at 0°C. After 10 min the mixture

was added to 200 mg of support **2** suspended with a mixture of 0.26 mmol of **5** (or **6**) in 700  $\mu$ L of anhydrous DCM. After 5 hours, at room temperature, the support was exhaustively washed with DCM, pyridine and then treated with a solution of pyridine/acetic anhydride (2.0 mL, 9:1 v/v) for 30 min at room temperature. Finally the support was washed with DCM, Et<sub>2</sub>O and dried under reduced pressure. Incorporation yields of nucleoside **3** and **4** onto **2** were in the range of 65-80% (0.029-0.035 meq/g), determined by quantitative DMT test performed on dried and weighed samples of the obtained support **5** and **6**.

The treatment of 50 mg of support **6** with 1.5 mL of solution of 0.4 M of NaOH in methanol/water (4:1, v/v) for 17h at room temperature followed by neutralization of the alkaline solution with a 2 M solution of triethylammonium acetate, furnished product **7** (0.5 mg, after RP HPLC purification).

**7**= <sup>1</sup>H-NMR (D<sub>2</sub>O); 8.02, 7.50, 7.48, 7.42, 7.26, 7.18, 7.06, (1H each, acridine protons and H-6 thymine; 5.56 (1H, m, H-1'); 4.60 (1H, m, H-3'); 3.90 (3H, s, CH<sub>3</sub>-O); 3.90-3.70 (5H, overlapped signals, CH<sub>2</sub>-NH, CH<sub>2</sub>-N(1), H-4'); 3.60 (2H, m, CH<sub>2</sub>OH); 3.48 (2H, m, H<sub>2</sub>-5'); 2.25 (2H, m, H<sub>2</sub>-2'); 1.61 (3H, s, CH<sub>3</sub>-5); 1.85–1.40 (7H, m's, 3 CH<sub>2</sub> and CH) UV-VIS (CH<sub>3</sub>OH)  $\lambda_{\max}$ : 265 nm and 422 nm; ESI-MS: m/z: 613 (M+H<sup>+</sup>).

### Oligonucleotide synthesis; products **9**, **10**, **12**

Symmetric oligomer elongation. 50 mg (0.0016 meq) of support **5** were used for each synthesis in the automated DNA synthesizer following standard phosphoramidite chemistry, using 45 mg/ml solution of 3'-phosphoroamidite building block in anhydrous CH<sub>3</sub>CN. Commercially N(4)-acetyl protected cytidine 3'-phosphoramidite building block was used to avoid deamination when sodium hydroxide is used for final cleavage and deprotection. Triemer **9** and ON **10** were synthesized performing respectively one or eight coupling cycles with appropriate nucleotide unit. Coupling yields resulted always >98% (by DMT test). After completion of the desired ON sequence and final DMT removal, the support was treated with 1.0 mL of 0.4 M solution of NaOH in methanol/water (4:1 v/v) for 17h at room temperature. The filtered solution and washings, neutralized with 5.0 mL of a solution 2 M triethylammonium acetate, were concentrated under reduced pressure, redissolved in water and purified by HPLC. Crude **9** was purified on a RP C<sub>18</sub> column (4.6x250mm) eluted with a linear gradient of MeOH in TEAB buffer (pH 7.0, 0 to 10% in 50 min, flow 0.8 mL/min), retention time 35.5 min. The final product (45 units OD<sub>260</sub>, from 50 mg of **5**) was lyophilized and characterized by spectroscopic data. Crude **10** was purified on a Nucleogel SAX column (Machery-Nagel, 1000-8/46); buffer A: 20mM KH<sub>2</sub>PO<sub>4</sub> aq. solution, pH 7.0, containing 20% (v/v) CH<sub>3</sub>CN; buffer B: 1.0 M KCl, 20 mM KH<sub>2</sub>PO<sub>4</sub> aq. solution, pH 7.0, containing 20% (v/v) CH<sub>3</sub>CN; using a linear gradient from 0 to 100% B in 20 min; flow rate 1.0 mL/min. The collected peak at retention time 18.8 min, desalted on Sep-Pak column (C18), furnished 28 OD<sub>260</sub> units of pure **10**.

**9**= <sup>1</sup>H-NMR (D<sub>2</sub>O); significant protons at  $\delta$ : 7.95, 7.90, 7.85, 7.45, 7.40, 7.32, 7.20, 7.15, 7.02 (9H, partially overlapped, acridine protons and 3H-6); 6.32-6.40 (3H, m's, 3H-1'); 1.65, 1.68, 1.70 (3H each, s's, 3CH<sub>3</sub>-5); <sup>31</sup>P NMR (D<sub>2</sub>O)  $\delta$ : 1.55 and 1.38; UV-VIS (H<sub>2</sub>O)  $\lambda_{\max}$  = 265 and 422 nm. ESI-MS m/z: 1221 (M+H)<sup>+</sup> 1243 (M+Na)<sup>+</sup>.

**10**= Maldi TOF-MS (positive mode): calculated mass 5357.03; found: 5358.83 (M+H), 5380.86 (M+Na).

Asymmetric oligomer elongation. 50 mg (0.0015 meq) of support **6** was used for the synthesis of oligomer **12** following standard phosphoramidite chemistry as described for product **10**. After the elongation of the first ON tract, the final DMT group was removed and the resulting 5'-OH function capped by reaction with 1 mL of pyridine/Ac<sub>2</sub>O solution (4:1, v/v) leading to **11**. The



removal of Fmoc protecting group was achieved by piridine/Et<sub>3</sub>N treatment (1.5h, 95:5, v/v). The resulting 5'-OH support was successively coupled with required 3'-phosphoramidite derivatives to complete the desired ON sequence, following standard automated procedure. Crude **12**, was detached from the support, deprotected and purified as described for **10**. The collected peak at retention time 20.5 min, after desalting furnished 24 OD<sub>260</sub> units of pure **12**. After HPLC purification 24 OD<sub>260</sub> units of **12** were obtained. Maldi TOF-MS (positive mode): calculated mass 5342.03; found: 5343.65 (M+H), 5383.06 (M+K).

#### UV thermal denaturation experiments

The concentrations of the synthesized oligomers were determined spectrophotometrically at  $\lambda = 260$  nm at 80 °C, using the molar extinction coefficient calculated for unstacked oligonucleotides using the following extinction coefficients: 11700 (G); 8800 (T) cm<sup>-1</sup>M<sup>-1</sup>. Aqueous solution of 5 mM NaH<sub>2</sub>PO<sub>4</sub>, 140 mM KCl and 5 mM MgCl<sub>2</sub> (at pH 5.5, 6.0 and 6.6) was used for the melting experiments. Melting curves were recorded using a concentration of approximately 1  $\mu$ M of single strand in 1 mL of the tested solution in teflon sealed quartz cuvettes of 1 cm optical path length. The resulting solutions were then allowed to heat at 90°C for 5 min, then slowly cooled and kept at 20°C for 15 h. After thermal equilibration at 20 °C UV absorption at  $\lambda = 260$  was monitored as function of the temperature, increased at a rate of 0.5°C/min. The melting temperatures were determined as the maximum of the first derivative of the absorbance vs. temperature plots.

#### Circular dichroism.

CD spectra were registered in the same buffer used for UV melting experiments at 25°C in a 0.1 cm path length cuvette. The wavelength was varied from 200 to 340 nm at 5 nm min<sup>-1</sup>, and the spectra recorded with a response of 16s, at 2.0 nm bandwidth and normalized by subtraction of the background scan with buffer. The temperature was kept constant at 25°C with a thermoelectrically controlled cell holder (JASCO PTC-348).

### References

1. Lane, A.; Jenkins, T.C.; *Curr.Org.Chem*, **2001**, 5, 845-869
2. Arthanari, H.; Bolton, P.H.; *Chem. Biol.*, **2001**, 8, 221-230
3. Jenkins, T.C.; *Curr.Med.Chem*, **2000**, 7, 99-115
4. Sen, D.; Gilbert, W.; *Nature*, **1988**, 334, 364-366.
5. Vasquez, K.M.; Narayanan, L.; *Science*, **2000**, 290, 530-533
6. Barre, F.K. ; Helene, C.; *Proc.Natl.Acad.Sci. USA*, **2000**, 97, 3084-3088
7. Fenseld, G. ; Davies, D.R.; *J. Am. Chem. Soc.*, **1957**, 79, 2023.
8. Morgan, A.R. ; Wells, R.D.; *J. Mol. Biol.*, **1968**, 37, 63-80.
9. Vasquez, K.M. ; Weeks, D.L. ; *J. Biol. Chem.*, **2001**, 276, 38536-38541
10. Praseuth, D. ; Helene, C.; *Biochem. Biophys. Acta.*, **1999**, 181-206
11. Winters, T.A.; *Curr. Opin. Mol. Ther.*, **2000**, 2, 670-681
12. McGuffie, E.M. ; Catapano, C.V. ; *Nucleic Acids Research*, **2002**, 30, 2701-2709
13. Scaria, V. ; Will, S. ; *J. Biol. Chem.*, **1995**, 31, 7295-7303
14. Thuong, N.T. ; Helene, C. ; *Angew.Chem.Int.Ed.Engl.*, **1993**, 32, 666-690
15. Sun, J.S. ; Helene, C. et al.; *C. R. Acad. Sci.III*, **1991**, 20, 585-590
16. Radhakrishnan, I. ; Patel, D.J. et al.; *J. Mol. Biol.*, **1991**, 221, 1403-141
18. Hoogsteen, K. ; *Acta Crystallogr.*, **1963**, 16, 907-916
19. Voet, D. ; Rich, A. et al.; *Prog. Nucleic Acid Research Mol.Biol.*, **1970**, 16, 183-265

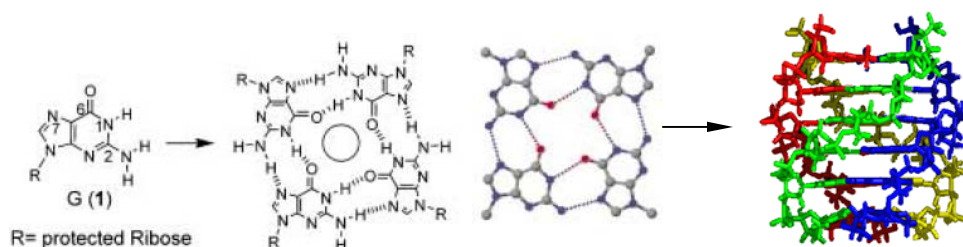
20. Saenger, W.; *Principle of nucleic acid structure*, Springer-Verlag, **1984**
21. Abrescia, N.G. ; Gonzalez, C. et al.; *Biochemistry*, **2004**, 43, 4092-4100
22. Aishima, J.; Gitti, R.K. et al.; *Nucleic Acid Research*, **2002**, 30, 5244-5252
23. Abrescia, N.G. ; Yhompson, A.; *Proc.Natl.Acad.Sci.USA*, 2002, 99, 2806-2811
24. Zagryadskaya, E.I. ; Doyon, F.R. et al.; *Nucleic Acid Research*, **2003**, 31, 3946-3953
25. Sharma, S.; Doherty, R.M.; *Curr. Med. Chem. Anticancer Agents*, **2005**, 5, 183-199
26. Bacolla, A.; Wells, R.D.; *J.Biol.Chem.*, **2004**, 279, 47411-47414
27. Ciotti, P.; Musso, M. et al.; *Eur.J. Biochem.*, **2001**, 268, 225-234
28. Sugimoto, N.; Wu, P. et al.; *Biochemistry*, **2001**, 40, 9396-9405
29. Sugimoto, N.; Wu, P. et al.; *Biochemistry*, **2001**, 40, 9396-9405
30. Nakano, S.; Fujimoto, M. et al.; *Nucleic Acids Research*, **1999**, 27, 2957-2965
31. Shiman, R. ; Draper, D.E et al.; *J. Mol. Biol.* , **2000**, 302, 79-91
32. Oh, D.H.; Hanawalt, P.C.; *Nucleic Acids Research*, **1999**, 27, 4734-4742
33. Vasquez, K.M.; Wang, G.; *Science*, **2000**, 290, 530-533
34. Faruqi, A.F.; Egholm, M. et al.; *Proc. Natl. Acad. Sci. USA*, **1998**, 95, 1398-1403
35. Faruqi, A.F.; Glazer, P.M. et al.; *Mol. Cell. Biology*, **2000**, 20, 990-1000
36. Vasquez, K.M.; Glazer, P.M. et al.; *Science*, **2000**, 290, 530-533
37. Chan, P.P.; Glazer, P.M. et al.; *J. Biol. Chem.*, **1999**, 274, 11541-11548
38. Gamper, H.B.; Meyer, R.B. et al.; *J. Am. Chem. Soc.*, **1999**, 120, 2182-2183
39. Francois, J.C.; Helene, C. et al.; *Proc. Natl. Acad. Sci. USA*, **1989**, 86, 9702-9706
40. Lee, H.C.; Murakami, A.; Miller, P.S.; *Biochemistry*, **1988**, 27, 3197-3203
41. Arimondo, P.; Bouturine, A.; *J.Biol.Chem.*, **2002**, 277, 3132-3140
42. Galeone, A.; Mayol, L. et al.; *Bioorg.Med.Chem.*, **2001**, 11, 383-386
43. Napoli, S.; Negri, U. et al.; *Nucleic Acids Research*, **2006**, 34, 734-736
44. Cuenoud, B.; Casset, F.; *Angew Chem. Int.Ed.Engl.*, **1998**, 37, 1288-1291
45. Carlomagno, T.; Blommers, M.J.J.; *J.Am.Chem.Soc.*, **2001**, 123, 7364-7370
46. Atsumi, N.; Matsuda, A.; *Bioorg.Chem.Med.*, **2002**, 10, 2933-2939
47. Cassidy, R.A.; Puri, N.; *Nucleic Acids Research*, **2003**, 31, 4099-4108
48. Buchini, S.; Leumann, C.J.; *Curr. Op. Chem. Biol.*, **2003**, 7, 717-726
49. Bernal-Mendez, E.; Leumann, C.J., *Nucleosides, Nucleotides Nucleic acids*, **2001**, 276, 35320
50. Obika, S.; Uneda, T. et al.; *Bioorg.Med.Chem.*, **2001**, 9, 1001-1011
51. Koizumi, M.; Morita, K. et al.; *Nucleic Acids Res.*, **2003**, 31, 3267-3273
52. Brodin, P. Sun, J.F. et al.; *Nucleic Acids Res.*, **1999**, 27, 3029-3034
53. Ueno, Y.; Mikawa, M. et al.; *Bioconjugate Chem*, **2001**, 12, 635-642
54. Filichev, V.V.; Nielsen, M.C. et al.; *Ang.Chem.Int.Ed*, **2006**, 45, 5311-5315
55. Mitsunobu, *Synthesis*, **1981**, 1-28
56. de Champdore' M., De Napoli, L. et al.; *Chem. Commun.* **2001**, 2598-2599.
57. Xodo, L. M.; Manzini G.; Quadrifoglio F. et al.; *Nucleic Acids Res.* **1990**, 18, 3557-3564.
58. Antao, V. P.; Grey, D. M ; Ratliff, R. L.; *Nucleic Acids Res.* **1988**, 16, 719-783.

## Chapter 3

### ***“Synthesis and characterization of new kind ON analogues able to form stable monomolecular quadruple helices DNA complexes: Tetra-End-Linked Oligonucleotides (TEL-ON)”***

#### ***Introduction***

In 1990 Guschlbauer, Chantot and Thiele published the review article “Four-Stranded Nucleic Acid Structures 25 Years Later: From Guanosine Gels to Telomere DNA”<sup>[1]</sup>, in which they highlighted the emerging importance of the G-quartet, a hydrogen-bonded ionophore<sup>[2]</sup> first identified in 1962 (Figure 3.1). Renewed attention to the G-quartet, in the late 1980s, was generated by intriguing proposals that the motif, when formed in DNA, might be biologically relevant<sup>[3,4]</sup>. Today, thousands of reports on some aspect of G-quartet structure or function have appeared, including some excellent reviews<sup>[5-10]</sup>. A G-quartet is a planar association of four guanines held together by eight Hoogsteen hydrogen bonds (figure 3.1); G quadruplexes result from the hydrophobic stacking of several quartets (figure 3.1)<sup>[4]</sup>. A cation (typically Na<sup>+</sup> or K<sup>+</sup>) is located between two quartets, the first forming cation–dipole interactions with four guanine, the second interacts with eight guanines, reducing the repulsion of the 2x4 central oxygen atoms, enhancing hydrogen bond strength and stabilizing quartet stacking.



**Figure 3.1:** G-quadruplex DNA

G-quartets may have applications in areas ranging from supramolecular chemistry to medicinal chemistry. Several reports suggest that DNA may be used as a building block for novel nano-sized objects<sup>[6]</sup>. Quadruplex DNA is an excellent module for the design of devices for nano-technology, because of its extreme rigidity and self recognition properties. G-quadruplexes are also likely to form higher-order structures such as synapsable DNA<sup>[8]</sup> or G-wires<sup>[9-10]</sup>. DNA nanodevices based on quadruplex–duplex interconversion<sup>[8]</sup> or biosensors<sup>[9]</sup>, have been also described in literature. The self-assembly of G-rich sequences can constitute liquid crystals<sup>[10]</sup> and may serve as the scaffold for artificial ion channels or receptors<sup>[11-12]</sup>. Quadruplex-prone segments may also be found in biologically significant regions such as telomeres<sup>[12-16]</sup> or oncogene promoter regions<sup>[17]</sup>, and a number of proteins or small molecules bind to G-quadruplexes<sup>[18-19]</sup>. In this contest, G-quadruplex DNA structures may provide a multitude of potential targets for drug design<sup>[20]</sup>.

### 3.1. Structural polymorphism of G-quadruplex DNA

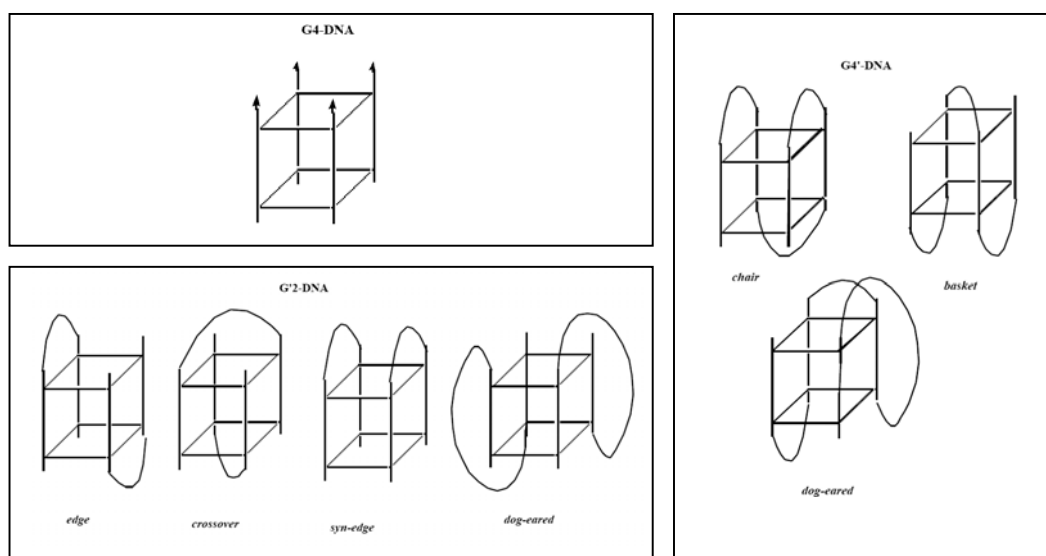
Numerous reviews have appeared recently discussing the structural characterization and polymorphism of G-quadruplex DNA<sup>[9]</sup>. It is now well established that certain DNA and RNA sequences, possessing a preponderance of guanosine residues, can form unusual structures. These structures exhibit unusual mobility under gel electrophoresis conditions, protection of the guanosine N7 from methylation with dimethyl sulfate, a characteristic CD spectrum, a signature cation-dependent stability to thermal denaturation, and distinctive imino resonances in the <sup>1</sup>H-NMR spectrum<sup>[1-10]</sup>. Early observations related to these structures came from the behaviour of guanosine nucleotides, which form gels in a cation-dependent manner. Fibre diffraction studies of the gels formed by guanosine-3'-monophosphate and poly(G) established the fundamental building block of these unusual structures, the G-tetrad, also called the G-quartet (figure 3.1). The G-tetrad consists of a planar arrangement of four guanine bases associated through a cyclic array of Hoogsteen hydrogen bonds in which each guanine base both accepts and donates two hydrogen bonds. The resulting square planar array is unique due to the "hole" that is created in the center. The G-tetrads form four-stranded helical structures by stacking interactions. In the case of the poly(G), the average distance separating the stacked G-tetrads is 3.4 Å. The tetrads are not stacked linearly, but adopt a right-handed helix. When one or more tetrads are stacked in this way, a cylindrical central cavity is produced. This cavity, lined with the guanine O-6 carbonyl oxygens, forms a specific binding site for metal ions. As deduced from the ability of particular metal ions to induce gel formation in guanosine nucleotides, the binding of metal ions follows the order  $\text{Sr}^{2+} > \text{K}^+ > \text{Rb}^+ \sim \text{Ba}^{2+} > \text{NH}_4^+ > \text{Ca}^{2+} > \text{Na}^+ > \text{Mg}^{2+} \sim \text{Cs}^+ > \text{Li}^+$ <sup>[9]</sup>. This contrasts sharply with duplex DNA, in which there is only a slight preference in metal ion binding in the series  $\text{Cs}^+ > \text{K}^+ > \text{Li}^+ > \text{Na}^+$ .

#### 3.1.1. Topological classification of G-Quadruplex DNA structures

A wide array of topologies and strand stoichiometries are compatible with the stacking of G-tetrads in helical structures. Hardin<sup>[21]</sup>, expanding on the naming convention suggested by Cech<sup>[22]</sup>, has proposed the class name G-DNA for this large class of structures; however, current usage appears to have settled on the class name G-quadruplex DNA. The structure formed from poly(G) consists of four strands, each contributing one G-residue to each G-tetrad. In this arrangement, the four strands are parallel, and there is a four-fold symmetry associated with the structure. This type of intermolecular, parallel-stranded structure has been referred to as G-quadruplex, G-tetraplex, or simply G4-DNA. In accord with the naming convention of Sen and Gilbert<sup>[23]</sup>, these four-stranded, parallel structures are referred to as G4-DNA (figure 3.2).

Alternatively, G-quadruplex DNA may form from two separate DNA strands, each of which contributes two guanosine residues to each G-tetrad. These bimolecular G-quadruplex DNA structures have been referred to as hairpin dimers. In keeping with the naming formalism of Sen and Gilbert<sup>[23]</sup>, this class of structures are referred to as G'2-DNA. There are a number of strand topologies that might be involved in the formation of G'2-DNA, but to date only four of these have been observed or proposed. As Bolton has noted, there are two distinct forms of G'2-DNA in which the loops formed by each of the two DNA strands are on opposite faces of the G-tetrad core<sup>[24]</sup>. In the edge form (figure 3.2), the loops connect adjacent DNA strands. A distinct crossover form is characterized by loops that connect strands that are diagonally related. A G'2-DNA topology in which the loops connecting DNA strands lie on the same face

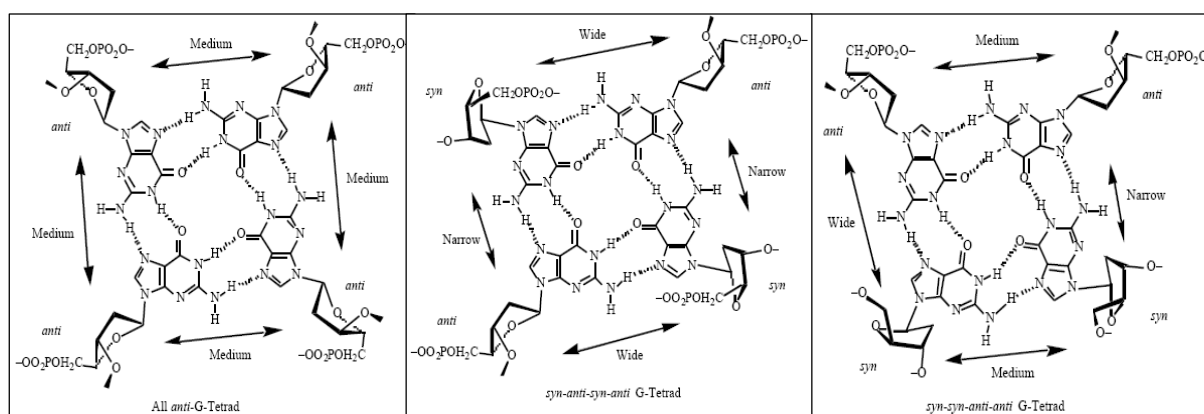
of the G-tetrad stack has been proposed by Sen and Gilbert (*syn-edge* G'-DNA, figure 3.2)<sup>[25]</sup>. An alternative topology for a bimolecular G-quadruplex DNA structure is a "dog-eared" structure (figure 3.2). This topology requires that the DNA loops run diagonally across the faces of the G-quadruplex that are perpendicular to the G-tetrads. Although this topology has not yet been observed in G'-DNA, a recently reported intramolecular G-quadruplex DNA has such a dog-eared" loop<sup>[26]</sup>. In addition to the G'-DNA topologies mentioned above, there are a variety of hypothetical topologies involving alternate strand distribution (i.e. three guanines in each G-tetrad from one strand and the fourth guanine from a second strand) or orientation (i.e. three parallel strands and one antiparallel strand, this latter recently observed<sup>[16]</sup> in human telomere structure forming an intramolecular 3+1 G-quadruplex scaffold in K<sup>+</sup>). It is likely that future research may uncover examples of these alternative topologies. Just as in the case of the bimolecular G-quadruplex DNA structures, there are multiple topologies that are possible in the formation of a unimolecular, or G4'-DNA. The chair form of G4'-DNA (figure 3.2) is characterized by loops connecting only adjacent strands<sup>[27]</sup>. This form is similar to the unimolecular form of G'-DNA proposed by Sundquist and Klug<sup>[28]</sup>. The basket form (Figure 3.2) has one loop that crosses the G-tetrad face diagonally<sup>[27]</sup>. A strikingly different folding topology is that reported by Wang and Patel<sup>[26]</sup> for the Na<sup>+</sup>-stabilized NMR structure of the *Tetrahymena* telomeric repeat d(T2G4)4. This sequence adopts a fold in which adjacent strands are connected by an edge loop on one face of the G-tetrad stack, a second, central lateral loop, and a terminal "dog-eared" loop running diagonally across one of the perpendicular faces of the G-tetrad stack. There are a multitude of alternative G4'-DNA fold topologies involving combinations of crossover, edge, and dog-eared loops; only time will tell if these topologies are energetically accessible to certain natural or engineered DNA sequences.



**Figure 3.2:** Different topological forms of termolecular (G4-DNA), bimolecular (G'-DNA), and unimolecular (G4'-DNA) G-quadruplex structures.

### 3.1.2. G-Quadruplex DNA groove widths

Related to the overall fold topology demonstrated by G'2-DNA and G4'-DNA are the glycosidic angles of the guanosine nucleotides forming the G-tetrads. In order for adjacent guanositones in antiparallel strands to form the hydrogen bonds necessary for G-tetrad formation, one of these guanositones must adopt a *syn* glycosidic conformation. Thus, for each topology listed above, there are alternative strand orientations that would result in distinct distribution of glycosidic conformations in the G-tetrads. For example, G'2-DNA having the edge topology can in principle adopt either a symmetrical form, in which the 5'-ends of the two strands are diagonally arranged about a common face, or an unsymmetrical form in which these two ends do not share any common faces. The symmetrical form requires that each G-tetrad have alternating *syn-anti-syn-anti* glycosidic conformations, whereas the unsymmetrical form requires *syn-syn-anti-anti* glycosidic conformations (figure 3.3).



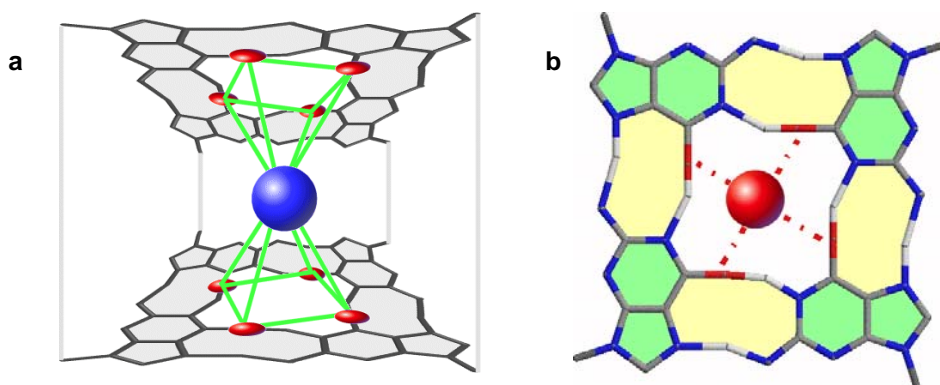
**Figure 3.3:** Variations in glycosidic torsional angles and their effects on groove width in G-quadruplex structures.

Experimentally, the symmetrical edge form is observed in the X-ray structure of  $[d(G4T4G4)]_2$ <sup>[29]</sup>. One consequence of this variation in glycosidic conformations, is the distinct nature of the grooves. The parallel G4-DNA G-tetrads, in which all G's adopt the *anti* glycosidic conformation, affords four widths of identical medium width. In cases where a guanosine in the *syn* conformation donates hydrogen bonds to a neighboring guanosine in the *anti* conformation, the groove formed between the two is extremely narrow, with a phosphate to phosphate distance as small as 7-9 Å<sup>[30]</sup>. In contrast, when the hydrogen bonding polarity between adjacent *syn*- and *anti*-guanosines is reversed, a very wide groove is formed<sup>[30]</sup>. An intermediate width groove results when adjacent guanositones adopt the same glycosidic conformation (figure 3.3)<sup>[30]</sup>. For example, in the structure of the symmetrical G'2-DNA having the edge topology, the G-tetrads are formed from guanositones of alternating *anti-syn-anti-syn* conformation, with each *syn*-guanosine donating hydrogen bonds to an adjacent *anti*-guanosine, and accepting hydrogen bonds from the other adjacent guanosine<sup>[29]</sup>. This results in a rectangular G-tetrad core with grooves of alternating wide-narrow-wide-narrow widths. The crossover G'2-DNA has guanositones that adopt the *syn-syn-anti-anti* conformations, which results in G-tetrads that adopt a parallelogram arrangement and the

formation of alternative wide, medium, narrow, medium width grooves between strands. Individual DNA strands in G-quadruplex structures is also a source of structural diversity, manifest by altered groove geometry. Both *syn-anti-syn-anti* and *syn-syn-anti-anti* G-tetrads typically stack upon each other such that consecutive G's in each DNA strand adopt an alternating 5'-3' *syn-anti* relationship. However, in G4'-DNA structures containing an odd number of stacked G-tetrads, unique *syn-syn* G-G steps are observed<sup>[29-31]</sup>. Through a combination of the altered base-edges and twist geometry associated with these *syn-syn* G-dinucleotide steps, the narrow grooves of these particular G4'-DNA structures provide a unique target for ligand design<sup>[26-31]</sup>. The ability of DNA minor groove ligands to distinguish the variable groove width engendered by sequence has been reported<sup>[32-33]</sup>. The specific array of alternating groove widths in G-quadruplex DNA structures should provide a means of selectively targeting specific topological and conformational forms of G-quadruplex DNA. For example, ligands which are able to span two or more G-quadruplex grooves may be able to distinguish one form from another based upon the specific groove dimensions and their specific orientation.

### 3.1.3. *G-Quadruplex DNA metal ion binding*

As has been observed for the association of individual guanosine nucleotides, G-quadruplex structures are stabilized by specific metal ion interactions<sup>[34]</sup>. The central cavity formed by stacked G-tetrads serves as a host to a variety of cations<sup>[35]</sup>. The metal cation preference for G4-DNA follows the trend  $K^+ > Ca^{2+} > Na^+ > Mg^{2+} > Li^+$  and  $K^+ > Rb^+ > Cs^+$ , in accord with the trend observed for guanosine nucleotide gel formation. In the case of  $K^+$ , the stacked G-tetrads in G-quadruplex DNA provide a bipyramidal antiprismatic coordination geometry (figure 3.4a), with each of eight carbonyl oxygens interacting equally with the cation.



**Figure 3.4:** Metal ion coordination geometry. a)  $K^+$ ; b)  $Na^+$ .

In the crystal structure of  $[d(G4T4G4)]_2$ , a region of weak electron density was found between planes of stacked G-tetrads and this was assigned to bound  $K^+$ <sup>[36]</sup>. In the high resolution crystal structure of  $[d(TG4T)]_4$ , two G4-DNA structures stack coaxially upon each other, forming an extended cylindrical central cavity<sup>[36]</sup>. Sodium ions are located between G-tetrad planes in the center of this cavity (figure 3.4b), but stack out of phase with the G-tetrad repeats, so as to bind in the terminal G-tetrad planes<sup>[36]</sup>. The selectivity of G-quadruplex DNA for  $K^+$  versus  $Na^+$  ion has been studied by Feigon and co-workers using  $^1H$ -NMR and the G'2-DNA formed by  $d(G3T4G3)$ , as a model system<sup>[37]</sup>. In this system, the G'2-DNA binds two  $K^+$  or  $Na^+$  ions, corresponding to one metal ion sandwiched between each pair of G-tetrads.

Using competition NMR experiments, Feigon and co-workers determined that there is only a modest difference in free energy of 1.7 kcal/mole, favoring the binding of  $K^+$  versus  $Na^+$  ions<sup>[37]</sup>. They suggest that this modest free energy difference is a result of the contributions of the relative free energies of hydration, which favors  $K^+$  binding. The effect of various cations on the kinetics of G-quadruplex DNA formation and the conformational polymorphism of G-quadruplex DNA has also been studied. Although original observations by Sen and Gilbert<sup>[38]</sup> indicated that  $K^+$  prevented the formation of G4-DNA from DNA oligonucleotides containing multiple runs of G's, this appears to be a kinetic, rather than thermodynamic phenomenon. G-quadruplex DNA structures can be trapped in metastable states in the presence of  $K^+$  ions, which stabilize and inhibit the interconversion of G-quadruplex DNA structures. Thomas<sup>[39]</sup>, Bolton<sup>[27]</sup>, and Hardin<sup>[40]</sup> have shown that the specific complexation of  $K^+$  by G-quadruplex DNA leads to a preference for the formation of G4-DNA structures under high  $K^+$  concentrations. Low  $K^+$  concentrations, or higher  $Na^+/K^+$  ratios, favour G4'-DNA structures. Whether a linear or a folded structure is formed is dependent on the exact sequence and number of repeats per strand, cations have a strong influence on their relative stabilities. It has been found that potassium ions preferably stabilize linear, four-stranded parallel quadruplexes whereas sodium ions stabilize folded forms. Cations have profound effects on loop geometry in G4'-DNA structures<sup>[27]</sup>. DNA oligonucleotides, with four runs of -GG- separated by two to four bases, can form chair G4'-DNA structures, but only in the presence of  $K^+$ <sup>[27]</sup>. When the -GG- runs are separated by four residues, a basket G4'-DNA structure can form in either  $Na^+$  or  $Na^+/K^+$ <sup>[27]</sup>. Sequences with runs of three or four G's separated by two to four bases can form basket G4'-DNA structures in the absence of  $K^+$ <sup>[27]</sup>. The presence of a purine in the loop can block both  $K^+$  binding and formation of the chair G4'-DNA structure. Modelling indicates that the requirement for  $K^+$  ion for chair formation is due to the ability of this ion to interact with the terminal G-tetrad and residues in the adjacent loop<sup>[27]</sup>. In addition to these cation-induced changes in G-quadruplex DNA topology, there is evidence for  $K^+$ -ion facilitated formation of G-quadruplex DNA structures from duplex DNA precursors<sup>[34]</sup>. Increasing  $K^+$  causes an increased cooperativity of G4-DNA thermal dissociation, despite the increased stabilization of G4-DNA at these higher  $K^+$  concentrations. Structural transitions of G-quadruplex DNA, induced by changes in univalent cation concentrations (i.e.  $K^+$  and  $Na^+$ ), may play an important role in DNA replication, organization, and function<sup>[39-40]</sup>. The transient changes in  $K^+$  concentration during the cell cycle may drive structural transitions in G-rich DNA. Low basal  $K^+$  concentrations (ca. 110 mM<sup>[41]</sup>) are thought to stabilize the G4'-DNA structures that may form from single-stranded G-rich sequences at the ends of the telomeres. As the cell cycle proceeds,  $K^+$  concentrations increase (ca. 130mM<sup>[42]</sup>) and this temporal elevation in  $K^+$  concentration initially favours the formation of G'-DNA structures, which may facilitate the alignment of chromosomal pairing in preparation for homologous recombination events. Further elevation in  $K^+$  concentration subsequently promotes the formation of G4-DNA structures that may be involved in the process of meiotic (or mitotic) telomere association<sup>[16]</sup>. An alternative view is that altered  $K^+/Na^+$  ratios in cancer cells may result in altered G-quadruplex DNA structure and function. A number of reports have documented altered  $K^+/Na^+$  ratios in tumour cells, although the magnitude and trend of change appear to be tumour type specific<sup>[43]</sup>. These cancer cell-specific  $K^+/Na^+$  ratios may effect conformational changes in G-quadruplex structures, resulting in altered transcription of particular genes and altered telomere function and maintenance.



### 3.2. Other Tetrads

In addition to the arrangement of four guanine bases, depicted in figure 3.1, there has been a variety of alternative tetrad arrangements of DNA bases observed or proposed in G-quadruplex structures. The inosine base also can adopt a tetrad structure, as has been observed in the gel formed by poly(I)<sup>[44]</sup>. Examples of stable assemblies containing combinations of stacked G•C•G•C<sup>[45]</sup>, G•T•G•T<sup>[46]</sup> and A•T•A•T tetrads are reported in the literature. Others have also demonstrated that a T•T•T•T tetrad can be accommodated when sandwiched between adjacent G-tetrads, although when placed at the ends of G-stacks the T-tetrads are largely disordered<sup>[47]</sup>. T-tetrad are less stable than G-counterpart, its geometry is stabilized through only four inter-strand hydrogen bonding that involve the O6 and N1 of each T-residue. The stability of A tetrads also appears to be context-dependent. Two possible geometries are envisaged in which the A-tetrad is stabilised through a putative interstrand hydrogen bonding between the 6-NH2 group and the N7, or 6-NH2 group and N1<sup>[48]</sup>.

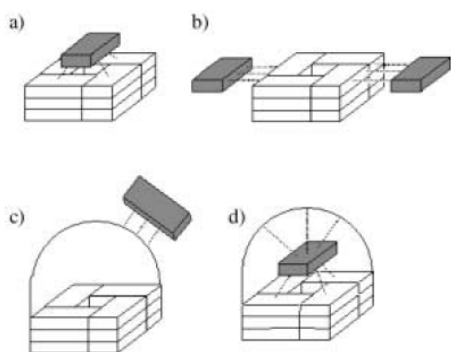
### 3.3. Biological relevance of G-quadruplex DNA

One of the strongest arguments in favour of the biological relevance of G-quadruplex DNA comes from the multitude of proteins that have been shown to bind these DNA structures. Moreover, many of these proteins bind only specific topological forms of G-quadruplex DNA. G-quadruplex forming elements can be found in important regions of human genome<sup>[1-9]</sup> in the telomeres, in the immunoglobulin switch regions, in the promoter regions of c-myc and other oncogenes, in the retinoblastoma susceptibility gene and upstream of the insulin gene. In principle, gene expression regulation can be obtained by modulation of a specific DNA binding protein, as a transcription factor, and directly intervening on the promoter of the recognized gene. Selectively interfering with these recognition processes, by exploiting small organic molecules specifically binding GROs (G-rich oligonucleotides), as the ones present in promoters of c-myc or K-ras oncogenes, which favour quadruplex formation, is a new and challenging therapeutic frontier. In this context, an interesting study on the promoter of c-myc oncogene recently appeared in the literature, which shows the direct evidence of *in vivo* existence of G-quadruplex structures<sup>[8]</sup>. DNA quadruplex structures are also thought to be involved, *in vivo*, in the inhibition of telomerase, a nucleoprotein devoted to maintain the length of chromosomes. The single stranded overhang of 5'-TTAGGG-3' repeats in the human telomere has been proposed to fold into an intramolecular G-quadruplex, and ionic conditions that stabilize G-quadruplexes inhibit telomerase. One approach to inhibit telomerase is to block the enzyme–substrate interaction. Whereas single-stranded DNA is a telomerase substrate, G-quadruplex DNA is not. Folding of telomeric DNA into G-quartet structures seems to influence the extent of telomere elongation *in vitro* and might therefore act as a negative regulator of elongation *in vivo*. A particularly appealing pharmacological approach is therefore based on the possibility to inhibit telomerase activity in tumour cells, where it results over-expressed, and to promote, as a consequence, their apoptosis by exploiting suitable drugs able to stabilize quadruplex structures in telomeres. Additional interest for quadruplex structures arises from the fact that they constitute the main scaffold of aptamers<sup>[13]</sup>, i.e. oligonucleotide-based molecules able to specifically interact with a given protein, thus modulating its function. Short DNA or RNA fragments are able to specifically bind proteins, as for example thrombin or HIV-integrase: therefore, these molecules cannot

be considered only as mere passive carriers of gene information, but as biomolecules directly involved in a complex series of cellular processes as well. On these grounds, synthetic GROs can represent a new class of pharmacologically interesting molecules, characterized by a high selectivity of action.

### 3.4. G-quadruplex ligands: molecular recognition

The wide array of G-quadruplex topologies, groove widths, loop conformations, and alternative DNA base-associations observed to date is probably just a small sampling of the structural diversity exemplified by G-quadruplex DNA. The unique features of G-quadruplex DNA: the association of up to four separate DNA strands, the hydrophobic stacking of large, planar hydrogen bonded DNA base quartets, and the specific interactions with metal ions, provide the basis for building up these diverse structures. Once formed, the G-tetrads of G-quadruplex DNA, can provide an "assembly floor" for the creation of unusual DNA base associations. The potential existence of G-quadruplex DNA *in vivo* opens up new vistas for the targeting of DNA sequences by molecular recognition of these structurally diverse DNA motifs. The structural uniqueness of the G-quadruplex can only aid in the realization of this ultimate goal in DNA sequence recognition. As shown in figure 3.5 and table 3.1, there are different ways in which molecules can interact with G-quadruplexes: by face recognition, edge recognition, loop recognition, or by simultaneous binding to the surface of the G-quartet and an adjacent loop or groove.



**Figure 3.5:** Schematic representation showing some of the ways that other molecules recognize a bimolecular G-quadruplex.

**Table 3.1:** Molecular recognition of G-quadruplexes

a) face recognition through stacking:

- nucleobase triads and tetrads
- aromatic molecules/telomerase inhibitors

b) edge recognition through hydrogen bonding:

- hydrogen bonding of other nucleobases
- protein recognition: edge interaction with amino acids
- water-nucleobase interactions

c) loop recognition through hydrogen bonding and electrostatic interactions:

- protein interactions: electrostatic phosphate -ammonium interactions
- small molecules and water

d) groove recognition through hydrogen bonding and electrostatic interactions

- water interactions with the hydration spine
- side-chain interactions with stacked aromatic molecules

Within the context of nucleic acids, there are two major ways for adjacent nucleobases to recognize a G-quartet: either by stacking on its surface or by hydrogen bonding to an exposed edge. Molecular recognition of the G-Quartet Edges Hydrogen-bond donor (N2-H and C8-H) and acceptor (N3) atoms on the exposed edges of the G-quartet are also molecular recognition sites. These sites, which are located along the floor of a groove running the length of the G-quadruplex, can bind water, ions, amino acid side chains, and other

nucleobases. Furthermore, neighboring loops in bimolecular and unimolecular G-quadruplexes are also molecular recognition sites, particularly for proteins and small molecules.

#### 3.4.1. *Molecular Recognition of G-Quadruplexes by Proteins*

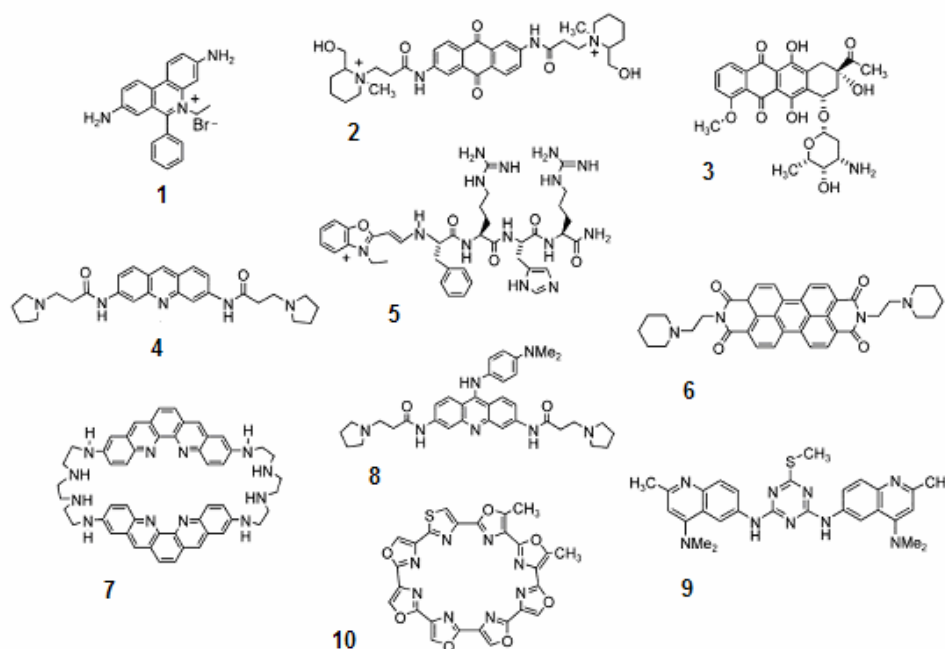
DNA and RNA oligonucleotides called aptamers have been selected by in vitro selection to bind molecular targets<sup>[49]</sup>. These targets are often proteins. One well-known aptamer is the thrombin binding aptamer (TBA), a DNA 15 mer that inhibits clotting. Crystallographic<sup>[50]</sup> and NMR<sup>[51]</sup> spectroscopic studies have shown that TBA forms a unimolecular G-quadruplex with two G-quartets and three loops. The structure and biological activity of TBA are K<sup>+</sup>-dependent. NMR spectroscopic, calorimetric, and ESI-MS studies have shown, however, that TBA is even more stable with divalent Pb<sup>2+</sup>, Sr<sup>2+</sup>, and Ba<sup>2+</sup> ions<sup>[52]</sup>. Related DNA aptamers are inhibitors of human HIV integrase with IC<sub>50</sub> values in the nm region<sup>[53]</sup> and adopt a unimolecular G-quadruplex similar to TBA<sup>[54]</sup>. An intermolecular G-quadruplex aptamer that bound the V3 loop of HIV reverse transcriptase was formed by a 17-mer phosphorothioate oligonucleotide<sup>[55]</sup>. Like many G-quadruplexes, this phosphorothioate aptamer was extremely stable and had a dissociation half-life of 60 days. Many other protein-binding aptamers have been proposed to form G-quadruplexes as part of their bioactive structure<sup>[56]</sup>. Although many proteins, including aptamers, oncogenic factors<sup>[57]</sup>, antibodies<sup>[58]</sup>, and telomeric proteins<sup>[10]</sup>, all bind G-quadruplexes, there are few molecular details known about quadruplex–protein interactions. The crystal structure of TBA bound to thrombin showed some ion pairs between the phosphate groups in the loops of the aptamer and Lys and Arg side chains<sup>[59]</sup>. A crystal structure of the telomeric protein of *Oxytricha nova* complexed to d[G4T4G4]4 is particularly valuable. As in the DNA solution structures, the bimolecular quadruplex in this complex has diagonal loops. Most of the DNA–protein contacts take place with these loops, rather than with the quadruplex core. However, the major DNA–protein interactions include: 1) electrostatic interactions, wherein the surface formed by three symmetry-related proteins provides a deep, electropositive cavity to hold the folded DNA; 2) van der Waals interactions, such as where the aromatic rings of Tyr142 and Phe141 pack against the G-4 sugar; 3) water mediated hydrogen bonds, wherein one of the T4 loops forms an extensive network of water-mediated hydrogen bonds with the protein; 4) nucleobase–peptide packing interactions, wherein the T6 residue of the other loop contacts the protein using both van der Waals and H-bond interactions. This T6 nucleobase packs between a Leu side chain and the polarized Asp437–Gly438 peptide bond, while simultaneously hydrogen bonding with atoms of the protein side chain and main chain. In both cases where crystal structures are available, for the *Oxytricha* and for the thrombin systems<sup>[59]</sup>, there are few direct interactions between the protein and the G-quadruplex core. Instead, the G-quadruplex seems to function as a scaffold on which loops are displayed for molecular recognition. Certainly, the crystal structure of the unimolecular G-quadruplex formed from the human telomeric sequence d(TTAGGG) suggests that the extended loops provide ideal protein-binding sites. In time, as more structural information becomes available, we should learn if a preference to bind loops is general for G-quadruplex protein interactions<sup>[57]</sup>.

### 3.4.2. *G-Quadruplex Aptamers : Recognition by Small Molecules*

Small molecules can also generate DNA and RNA aptamers, and G-quartets often appear as part of the aptamer structure<sup>[60]</sup>. Porphyrins and G-quartets have similar surface areas, and many biophysical and biochemical studies have demonstrated stacking interactions between porphyrins and G-quadruplexes<sup>[61]</sup>. Sen and co-workers also identified DNA aptamers able to catalyze reactions<sup>[62]</sup>. Some DNA–hemin complexes had enhanced peroxidase activity relative to the heme cofactor alone. They concluded that the folded DNA activates the bound heme<sup>[62]</sup>. A G-quadruplex, with a heme intercalated between G-quartets was proposed to explain the enhanced peroxidase activity of the aptamers. These studies underscore the potential of DNA to function as a catalyst, in addition to its information storage role. By using in vitro selection, Isalen et al. discovered Zn<sup>2+</sup> finger peptides that bind G-quadruplexes<sup>[63]</sup>. High affinity peptides (K<sub>a</sub>=25 nM) containing three helices bound the human telomeric sequence d(GGTAG). These Zn<sup>2+</sup> fingers were G-quadruplex-specific; duplex DNA with the same sequence was not bound. The Zn<sup>2+</sup> fingers contained high amounts of certain amino acids, and the authors speculated that Glu/Asp side chains might form hydrogen bond with the N2-amino groups of the guanosine, while His groups could stack on the G-quartet.

### 3.4.3. *Small Molecules that Stabilize DNA G-Quadruplexes: Telomerase Inhibitors*

Telomerase inhibition depends on shifting the equilibrium between the single-strand and quadruplex towards the folded form. One strategy is to stabilize the G-quadruplex form by binding high-affinity small molecules. A number of reviews describe well how drug design and screening efforts have identified compounds that bind G-quadruplexes and inhibit telomerase. As shown by the structural formulas **1–10** (figure 3.6), most telomerase inhibitors are aromatic compounds with electron-deficient rings suited for stacking with the electron rich G-quartet. Ethidium bromide (**1**), a well-known duplex intercalator, binds DNA quadruplexes<sup>[64]</sup>, but its surface area doesn't quite cover a G-quartet, and it was reasoned that larger aromatic compounds would be better inhibitors. The first report of a small molecule, anthraquinone derivative **2**, that binds G-quadruplex DNA and inhibits telomerase came in 1997<sup>[65]</sup>. Since then, hundreds of telomerase inhibitors have appeared. There is debate about how the aromatic inhibitors interact with G-quadruplex DNA. Some suggest that they intercalate between G-quartets, while others argue that the compounds end-stack on terminal G-quartets. The few known quadruplex–ligand structures, obtained from NMR spectroscopic<sup>[66]</sup>, fiber diffraction<sup>[67]</sup>, and single-crystal X-ray crystallographic analysis<sup>[67]</sup> indicate that telomerase inhibitors end-stack rather than intercalate. In their recent X-ray study of interactions between G-quadruplex and daunomycin (**3**) Clark et al. noted that intercalation of drugs between G-quartet layers should be energetically more costly than end stacking, as intercalation requires unstacking of the tetrad and unwinding of the helix. In addition, end stacking of aromatic drugs, as opposed to intercalation, would allow stabilizing cations to remain coordinated within the G-quadruplex core. Structure-based design and activity-based screening have both identified telomerase inhibitors. For example, Shafer and co-workers used computer modelling to screen “virtual” libraries for compounds that might interact with G-quadruplexes. These efforts identified a carbocyanine dye that binds bimolecular G-quadruplexes<sup>[69]</sup>. Molecular modeling studies suggested that addition of a properly positioned third side chain to an existing acridine drug **4** (to give **8**) would increase the affinity of the G-quadruplex.



**Figure 3.6:** High – affinity small molecules binding G-quadruplex.

They reasoned that additional contacts between the side chains and grooves would further stabilize the folded DNA, and thus lead to more telomerase inhibition. A comparison of the properties of acridines **4** and **8** showed that attachment of this third side chain improved the drug's affinity to bind with the G-quadruplex, its quadruplex–duplex selectivity, and its telomerase inhibitory activity. In another modeling approach, Fedorov, Hurley, and coworkers designed and synthesized the quadruplex interactive compound PIPER (**6**), which was a potent telomerase inhibitor<sup>[70]</sup>. The slow NMR exchange between DNA bound and free PIPER (**6**), along with ligand–DNA NOE interactions, allowed determination of the drug's location when bound to G-quadruplex DNA. The conclusion was that PIPER (**6**) end-stacked over a terminal G-quartet. Promising telomerase inhibitors have also come from combinatorial screening strategies. The research groups of Mergny and HQIRne used a FRET-based assay to identify triazine compounds, such as **9**, that induce a DNA oligonucleotide to fold into a G-quadruplex<sup>[71]</sup>. Some of these triazines were inhibitors of human telomerase at nanomolar concentrations. Such combinatorial screens can lead to active compounds that might not have been considered in structure based design. Finally, the most potent and selective telomerase inhibitor, telomestatin (**10**), is a natural product isolated from *Streptomyces anulatus* during an activity screen<sup>[72]</sup>. NMR and modeling studies indicate that telomestatin stacks on a G-quartet, which again suggests that telomerase inhibition derives from an ability to stabilize the G-quadruplex form of telomeric DNA<sup>[73]</sup>.

### 3.5. Guanosine self-assembly in materials science, biosensor design, and nanotechnology

Quadruplexes show also promise as components for nanowires, ion channels, and building blocks for directing the assembly of nanoscale components into sophisticated structures.

#### 3.5.1. *DNA Nanostructures: G-Wires, Frayed Wires, and Synapses*

Fiber diffraction studies on crystalline GMP and poly-(guanylic acid) had shown that these compounds self-assemble into rods of stacked G-quartets<sup>[74]</sup>. In the mid-1990s, the research groups of both Henderson and Sheardy further established that G-rich DNA formed nanometer assemblies<sup>[75]</sup>. They showed by using gel electrophoresis that d(G4T2G4), in the presence of K<sup>+</sup> and Mg<sup>2+</sup> ions formed high molecular weight assemblies that were well resolved from the smaller G-quadruplex [d(G4T2G4)]<sub>4</sub>. The polymers were extraordinarily stable; heating at 80°C or dissolving in 8M urea did not denature them. Marsh and Henderson coined the term “G-wires” to describe the continuous, parallel-stranded DNA superstructures formed when the 5'-end of one DNA-duplex with G-G pairs associates with the 3'-end of a similar duplex<sup>[76]</sup>. Marsh and Henderson proposed that G-wires could be useful in nanotechnology, nano-electronics and biosensor development.

#### 3.5.2. *Toward Synthetic Ion Channels*

G-quadruplexes can organize themselves to build transmembrane ion channels. NMR studies have shown that base pairs in DNA G-quadruplexes open slowly and G-quadruplex dissociation is often quite slow, taking days or weeks. These results suggest that ions move without disruption of the G-quartet, thus making the G-quadruplex analogous to an ion channel.

#### 3.5.3. *Formation of Biosensors and Nanomachines with G-Quadruplex DNA*

While DNA aptamers have potential as therapeutics and diagnostics, they also have applications in bioanalytical chemistry. Small molecules and proteins can be separated by using G-quadruplex DNA as stationary phases in chromatography or electrophoresis. DNA aptamers featuring an intramolecular G-quadruplex served as the stationary phase for the separation of the isomeric dipeptides Trp–Arg and Arg–Trp<sup>[77]</sup>. DNA G-quadruplex aptamers labeled with fluorescent dyes have also served as a prototype for biosensors<sup>[78]</sup>. In particular, FRET has been used to study the secondary structure of G-rich DNA oligonucleotides<sup>[79]</sup>. FRET is a distance-dependent method for detecting conformational changes over distances of 10–100 Å. FRET has also been used to monitor the formation of molecular “nanomotors” from single-stranded DNA<sup>[80]</sup>. Li and Tan used FRET to monitor in real time the conformational switching between a double-stranded duplex and a folded G-quadruplex caused an extension and shrinking motion<sup>[81]</sup>. In 1975 Aviram and Ratner proposed that organic molecules could conduct charge<sup>[82]</sup>. Today, with the miniaturization of semiconductor components, the potential of molecular electronics is even clearer<sup>[83]</sup>. Recent experiments with DNA suggest that biomolecular self-organization is a viable strategy for electron transport<sup>[84]</sup>. Guanine derivatives are attractive candidates for electron transport. Guanine has the lowest oxidation potential of the nucleobases and it forms a variety of ordered structures.

### 3.6. Future direction

Three decades after its identification in 5'-GMP gels, the G-quartet really gained visibility because of the biological implications of G-quadruplex DNA. Consequently, structural characterization of nucleic acid G-quadruplexes has accelerated over the past decade.

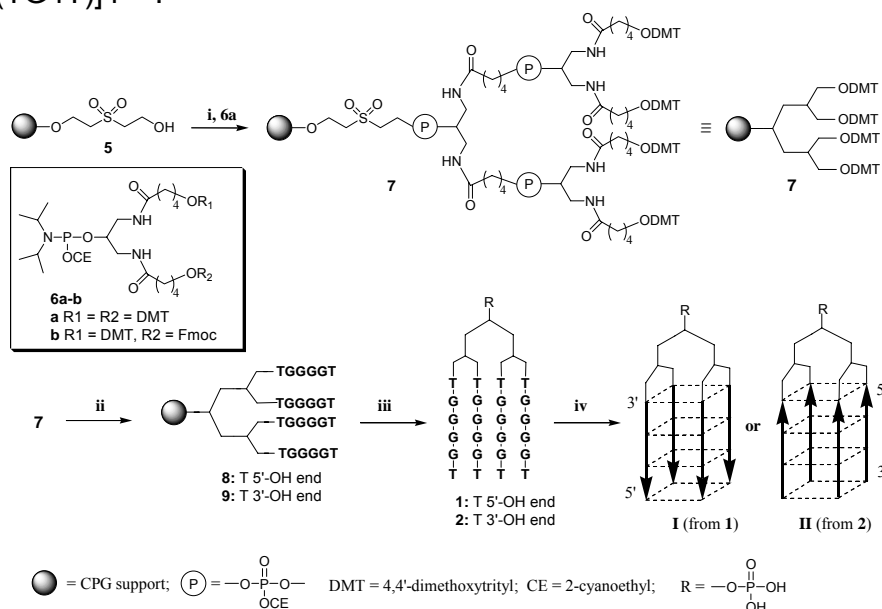
In 1992 the Protein Data Bank showed just two entries for G-quadruplexes. Ten years later, there are more than 35 deposited G-quadruplex structures, including 9 crystal structures. This data provides a wealth of information about the interactions that drive self-assembly of G-rich nucleic acids. More structures of G-quadruplex-protein and quadruplex-ligand complexes are needed in the future; indeed, the first three examples were published in the last three years<sup>[85]</sup>. Detailed information about protein and ligand complexes with G-quadruplexes will help us to better understand G-quadruplex molecular recognition and further enhance drug design. DNA quadruplex structures are very appealing from a pharmacological point of view, both as scaffolds of oligonucleotides (ONs) able to specifically inhibit proteins (aptamers) and in vivo, as targets in antitumour therapies. However, the pharmacological use of natural ON in vivo experiments is dramatically limited by factors concerning their stabilization, administration and bioavailability. In order to overcome these drawbacks, recently many research efforts have been devoted to the design and synthesis of ONs containing suitable chemical modifications, for example at 3'- or 5'-ends, in the backbones and others, in order to improve better applications in biomedicine and biochemistry.

In this context, my effort has been addressed to synthesize and study of a class of ON analogues forming quadruplex structures with potential biological properties. These studies would be a part of a more general and promising investigations aimed at using ON analogues as drugs able to selectively interfere with the normal functions of nucleic acids and specific cellular proteins, thus inducing a specific control of the gene expression. Potential applications not only as therapeutic aptamers, but also as chromatography supports, and molecular electronic devices may be promising and aim the synthesis and the structural studies of this new class of ON analogues, that I wish to propose.

### 3.7. The aim of the work

The potential of quadruplex DNA as targets for therapeutic intervention and as therapeutics is certain, and require, for its realization, more detailed structural informations about G-quadruplex selective targeting. Up to today, many of the studies have used ONs that may form multiple structures making it difficult to ascertain which quadruplex structure type, or types, were involved in the interaction examined. The use of ONs that form a known, single quadruplex structure will allow the interactions to be better characterized and may also indicate that the binding affinities are stronger than those deduced from results on the ONs that adopt multiple structures. In order to better understand G-quadruplex molecular recognition and further enhance drug design, simpler models of biologically relevant tetrads needed, in order to obtain high resolution data on drug-DNA interactions. Tetramolecular quadruplexes (also called G4-DNA) have been used at this purpose. First, these structures show simpler configuration. Second, the conformation of the stem of these quadruplex is very close to the central core of intramolecular parallel quadruplexes found for human telomeric repeats or other motifs. Third, several tetramolecular quadruplexes strongly interact with the HIV gp120 protein, and act as specific inhibitors of infections in vitro<sup>[86]</sup>. Unfortunately, the in vivo formation of intermolecular quadruplexes is very slow, requiring high ONs concentration.

These unfavourable kinetic and thermodynamic parameters could be disadvantageous in view of their potential therapeutic use. Starting from these observations, the first goal of my research work, in this context, was the synthesis of a new modified ON (TEL-ON, **1**, Scheme 3.1) capable of forming the monomolecular parallel G-quadruplex **I**, which thermal stability has been proved to be significantly higher than that observed for its tetramolecular counterpart  $[d(TG_4T)]_4$ <sup>[87]</sup>.



**Scheme 3.1:** i: two couplings with **6a**; ii: ODN synthesis with 3' or 5'-phosphoramidites; iii: detachment and deprotection with  $\text{NH}_4\text{OH}$  conc. 32% (7 h, 55 °C); iv: HPLC purification and annealing.

The main structural feature of this molecule is the presence of four ON strands whose 3'-ends are attached to a non-nucleotidic tetra-end-linker. The synthesis of the symmetric tetra-branched linker was carried out on polymeric support. This branched structure bears four primary alcoholic functions prone to ON chain assembly (for example TGGGT or TGGGXT, where X is a variable base). By choosing appropriate alcoholic protecting groups of the linker structure and the ON chain assembly procedure, TEL-ONs having predetermined strand orientation (parallel, antiparallel or both, **1-4**) or containing modified bases, were obtained<sup>[88]</sup>. Moreover, quadruplex complexes, containing less stable quartets such as T-tetrads<sup>[89]</sup>, were synthesized. Modulating the length and nature of linker was also performed in order both to generate quadruplexes of higher stability and to hinder the folding of the linker around the whole quadruplex structure, thus forcing the predetermination of quadruplex strands polarity upon the basis of the relative orientation of the ON strands into the TEL-ON molecules. Furthermore, the construction of new and stable monomolecular quadruplexes possessing unusual strand orientation could expand their employment in both the study and control of important biological processes as well as in DNA-based nanotechnologies. Several studies have investigated the effect of loops on quadruplex stability and typology. With regard to nucleotidic loops, the base composition, the length, and the orientation of the phosphodiester linkage have influenced quadruplex typology. Non-nucleotidic loops have also been investigated. In this case both the kinetics of the quadruplex formation and the complex stabilities were found to be affected by the nature of the linker. Tetra-End-Linked linker seems



to provide favourable kinetic and thermodynamic parameters. Preliminary studies confirm that TEL-ON are more stable than the corresponding tetramolecular counterpart. Briefly, my research work in quadruplex-DNA was employed and addressed on the synthesis and characterization of a new kind ON analogues able to form stable monomolecular quadruple helices DNA complexes (TEL-ON). Studies on the length of tetra-end-linked linker on the orientation of the ON strands aimed at obtaining stable quadruplex structures and studies on quadruplex complexes, containing less stable quartets such as T-tetrads, have been also performed and reported.

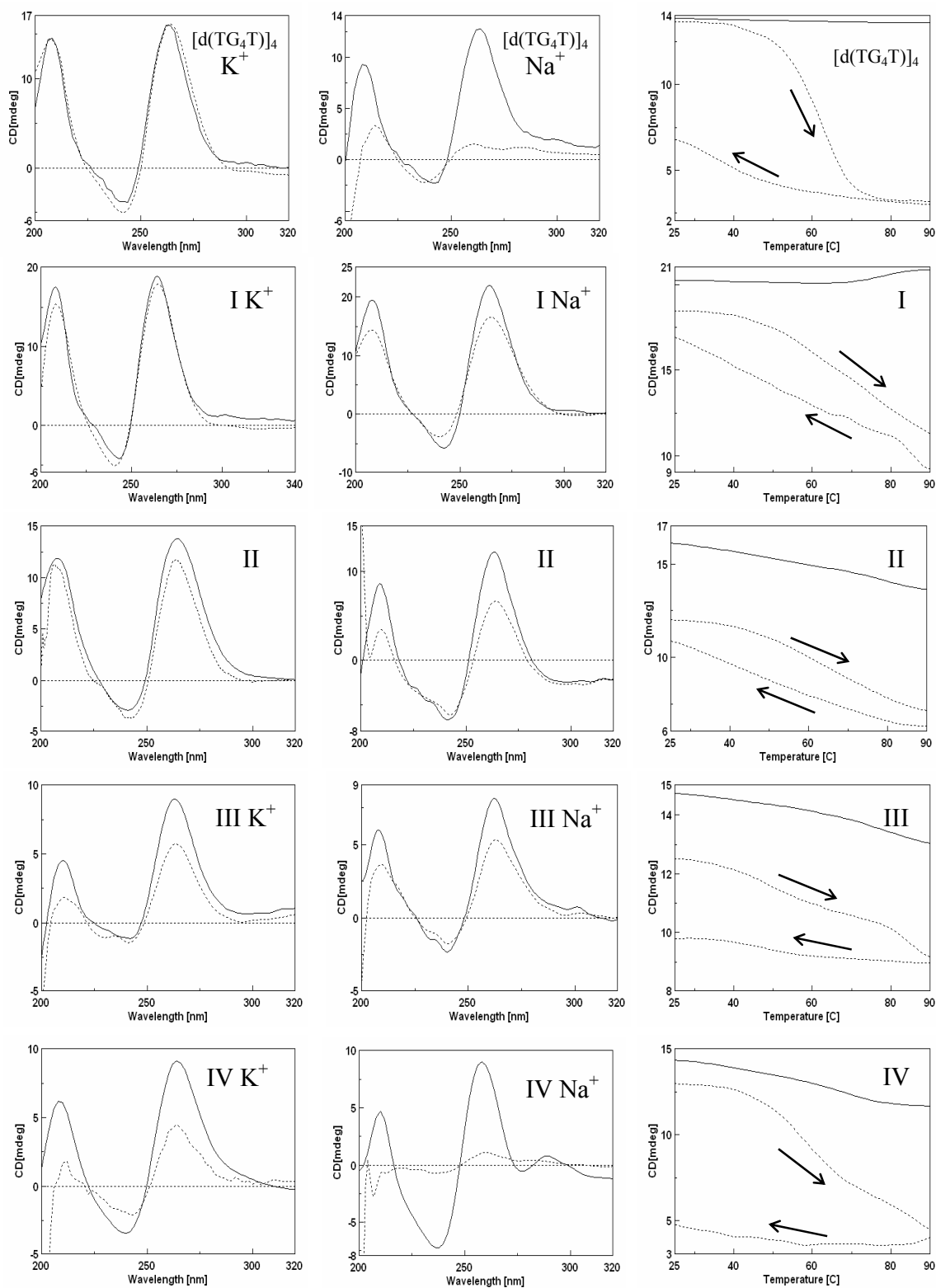
### 3.8. *Synthesis and characterization of a new kind ON analogues able to form stable monomolecular quadruple helices DNA complexes (TEL-ON)*

The first goal of my research on TEL-ON was the chemical synthesis and the structural characterization of TEL-[d(TG4T)]<sub>4</sub>, which represents the first example of a new class of ONs forming quadruplex. The sequence of the chains was chosen mainly because d(TG4T) forms a stable tetramolecular helix, already well characterized by thermodynamic measurement as well as NMR, X-ray crystallography and CD spectra<sup>[85]</sup>. This quadruplex possesses a fourfold symmetry with all strands parallel to each other and all nucleotides in an *anti* conformation. CD and NMR spectroscopies have been used to investigate the solution structure of TEL-[d(TG4T)]<sub>4</sub> in the presence of potassium or sodium ions in comparison with the natural counterpart [d(TG4T)]<sub>4</sub>.

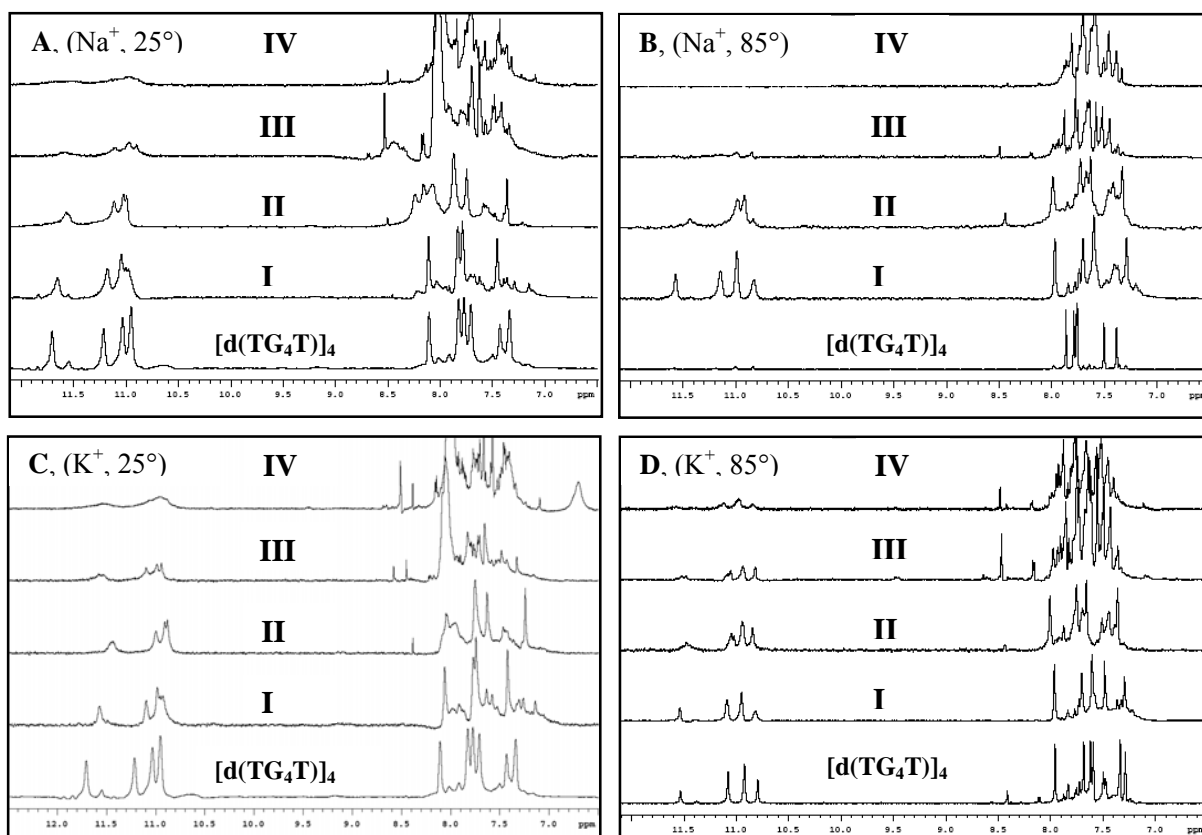
**Synthesis and Purification of TEL-ON (I).** The synthetic strategy employed starts from the commercially available Controlled Pore Glass (CPG) support **5** and the phosphoramidite derivative of the di-functional linker **6a**. On support **5** (0.048 meq/g), performing two coupling cycles with **6a** by an automatic DNA synthesizer, support **7** was obtained, that bears four symmetrical protected primary alcoholic functions, prone to ON-chain assembly (0.085 meq/g of OH functions, by DMT spectrophotometric test). ON chains were assembled on **7** utilizing 3'-phosphoramidite nucleotide units, thus obtaining the polymer bound ON **8** (coupling yields >98% per cycle). The presence of a  $\beta$ -sulfone phosphodiester function in the linker connecting the TEL-ON to the solid support allows the release of the phosphomonoester group via a  $\beta$ -elimination under alkaline treatment (conc. NH<sub>4</sub>OH, 7 h, 55° C), affording the TEL-ON **1**. Crude **1** was purified by HPLC on an anionic exchange column and the collected peaks were desalted on a RP18 column thus furnishing **1** (>95% pure, by HPLC and PAGE). In a typical synthetic cycle, starting from 100 mg (0.0085 meq) of support **5** 135 OD<sub>260</sub> units yield of pure **1** were obtained. The structure of **1** was confirmed by <sup>1</sup>H NMR data and by ESI mass spectrometry, which showed pattern peaks indicating a molecular weight at 8762.6 a.m.u. (calcd. 8763.7) (Scheme 4.1).

**CD and NMR studies of I and d(TG4T).** In order to demonstrate that **1** can adopt a G-quadruplex structure **I** and, in case, to estimate its stability, circular dichroism (CD) studies and CD thermal denaturation experiments were performed, in comparison with the tetramolecular G-quadruplex [d(TG4T)]<sub>4</sub>. CD spectra were registered at 20° C at quadruplex concentrations in the range 4.0–4.5·10<sup>-5</sup>M in K<sup>+</sup> or Na<sup>+</sup> buffers. Particularly, TEL-ON **I** and [d(TG4T)]<sub>4</sub> showed almost identical CD profiles in both saline conditions (figure 3.7, columns A, B), characterized by a positive and a negative Cotton effect at 264 and 249 nm, respectively, diagnostic of a G-quadruplex structure involving four parallel strands. CD thermal denaturation experiments (figure 3.7, column C) were performed monitoring the CD

values (mdeg) at 264 nm in the range 20–90°C with 1°C/min heating rate. In K<sup>+</sup> buffer both **I** and [d(TG4T)]<sub>4</sub> G-quadruplexes undergo irrelevant CD<sub>264</sub> variation up to 90°C, thus not showing the typical descendant sigmoidal curves indicative of a melting process. On the other hand, in Na<sup>+</sup> buffer, the increase of temperature leads to a significant reduction of CD<sub>264</sub> values for both complexes, thus indicating the fusion to occur. The complex [d(TG4T)]<sub>4</sub>, showed a defined sigmoidal curve with a derivatizable T<sub>m</sub> value (56°C at 4.5·10<sup>-5</sup>M concentration). In the case of **I** the melting curves registered at concentrations 1.0 · 10<sup>-5</sup>, 2.0 · 10<sup>-5</sup> and 4.5 · 10<sup>-5</sup>M did not show well-defined changes of convexity up to 90°C, thus suggesting that only partial fusions occur. Furthermore, it is noteworthy that the melting curve profiles of **I** are not concentration dependent thus indicating that the complex is mainly present as a monomolecular species. The whole of data suggests that **I**, in Na<sup>+</sup> buffer, is characterized by a higher stability than its tetramolecular counterpart [d(TG4T)]<sub>4</sub>, most likely due to the favourable entropy of the monomolecular complex. <sup>1</sup>H NMR studies (500 MHz) on both complexes confirmed the data obtained by CD. Spectra were recorded in Na<sup>+</sup> and K<sup>+</sup> buffers at 25, 65 and 85°C using pulse field gradient WATERGATE for H<sub>2</sub>O suppression. The spectrum of **I**, in Na<sup>+</sup> buffer at 25°C (figure 3.8A), shows the presence of four singlets in the range 10.5–12.0 ppm, two of them partially overlapping, confidently assigned to exchange protected imino protons involved in Hoogsteen N(1)H/O(6) hydrogen bonds of G-quartets also on the basis of the comparison with the corresponding region of the spectrum of [d(TG4T)]<sub>4</sub> (figure 3.8B), which displays a very similar peak profile. Even if a detailed <sup>1</sup>H-NMR study has not been yet accomplished, it is not unreasonable at this stage to hypothesize that **I** is characterized by a quadruplex structure containing four G-quartets with a pseudo fourfold symmetry, so that the imino protons of the four oligonucleotide chains resonate as they were magnetically equivalent. NMR spectra in Na<sup>+</sup> buffer at higher temperatures indicate that the above structure of **I** is preserved at least up to 85°C, whereas the tetramolecular complex [d(TG4T)]<sub>4</sub> is destroyed at this temperature, as indicated by the disappearance of the imino protons signals. The <sup>1</sup>H-NMR spectra in K<sup>+</sup> buffer, for both complexes, displayed quite similar signal patterns. However, it was not possible to observe any quenching of the imino protons up to 85°C also for [d(TG4T)]<sub>4</sub>. Even though TEL-[d(TG4T)]<sub>4</sub> is not the first monomolecular parallel quadruplex ever investigated, nevertheless, thanks to its unique structural features, it can be regarded as the prototype of a new class of quadruplex forming ONs, potentially useful for a number of interesting applications, considering the increasing number of ascertained or supposed biological roles in which quadruplex structures are involved. As a matter of fact, monomolecular quadruple helices formed by TEL-ONs are provided with more favourable thermodynamic and kinetic parameters than their tetramolecular counterparts. Therefore, TEL-ONs with suitable sequences and predetermined strand orientations could be used as aptamers or decoys with improved properties.



**Figure 3.7:** CD and CD melting spectra of I-IV and d[(TG4T)]4 in K<sup>+</sup> buffer (left column), in Na<sup>+</sup> buffer (middle column) at 25 °C (solid line) and 90 °C (dotted line). CD<sub>265</sub> melting profiles (right column) in K<sup>+</sup> (solid line) and Na<sup>+</sup> (dotted line).

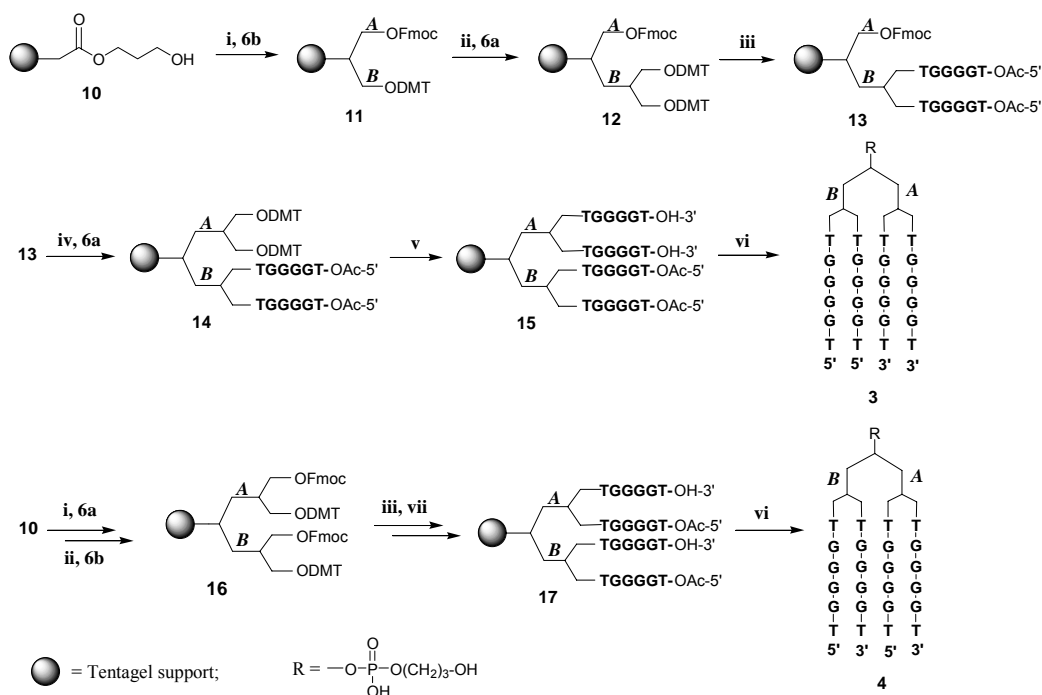


**Figure 3.8:**  $^1\text{H}$  NMR (500 MHz) spectra of I-IV and  $[\text{d}(\text{TG}_4\text{T})]_4$  in  $\text{Na}^+$  and  $\text{K}^+$  buffers. The spectra were registered in  $\text{H}_2\text{O}/\text{D}_2\text{O}$ , (9:1, v/v) solution at 0.5 mM of quadruplex at 25 and 85  $^\circ\text{C}$ . The aromatic and imino proton regions of the spectra are reported.

Starting from these perspectives, the syntheses of several parallel and/or antiparallel stranded TEL-ONs, also containing such types of unusual tetrads, have been performed. Two synthetic strategies have been used to obtain  $\text{d}(\text{TG}_4\text{T})$  ON strands which were linked with their 3' or 5'-ends to each of the four arms of the linker. The first approach allowed the synthesis of the TEL-ONs containing the four ON strands with a parallel orientation (**1**, **2** Scheme 3.1), while the latter synthetic pathway led to the synthesis of TEL-ONs each containing antiparallel ON pairs (**3**, **4**, Scheme 3.2). CD, CD melting, NMR spectroscopy, and molecular modeling have been used to investigate the solution structures of the resulting TEL-quadruplexes I-IV formed in the presence of sodium or potassium ions.

**Synthesis and Purification of TEL-ONs 1-4.** In the synthesis of the TEL-ONs **1-4** two hydroxyl functionalized solid supports (**5** and **10**) and the bifunctional linkers **6a,6b** have been used. The synthesis of **1** was just above described. **2** has been obtained following the same protocol employed for **1** (Scheme 3.1). **2** differs from **1** only in ON chain orientation, which synthesis require 5'-phosphoramidite nucleotides as building blocks. In both syntheses coupling yields were higher than 98% per cycle. As a further step of this work we synthesized the TEL-ONs **3** and **4** (Scheme 3.2). In **3** a pair of strands was linked to the primary linker branch A through the 5'-ends whereas the second pair is linked to the primary branch B through the 3'-ends. Instead **4** is characterized by both branches (A and B) carrying two strands linked through opposite ends. For the preparation of **3** and **4**, we adopted two

synthetic strategies both based upon bi-functional linker **6b** in which the two alcoholic functions, thanks to the orthogonal protection with Fmoc and DMT groups, allowed the sequential synthesis of the ON chains with opposite polarity. **10** was used as solid support since, when **5** was used, we observed the release of ON material from the support during the Fmoc deprotection step, due to the concomitant  $\beta$ -elimination at the sulfone group.

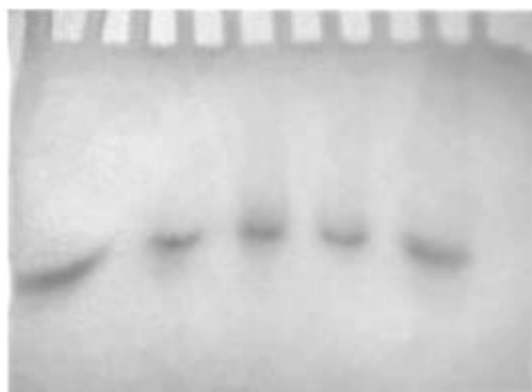


**Scheme 3.2:** i: coupling with linker 6; ii: DMT removal and coupling with linker 6; iii: ON synthesis with 3'-phosphoramidites; iv: Fmoc removal and coupling with linker 6a; v: ON synthesis with 5'-phosphoramidites; vi: detachment and deprotection with  $\text{NH}_4\text{OH}$  conc. 32% (7 h, 55°C) and HPLC purification; vii: Fmoc removal and ON synthesis with 5'-phosphoramidites

Therefore, we prepared the hydroxyl-functionalized support **10** bearing an ester function, stable under Fmoc deprotection conditions but cleavable by ammonia treatment, to release the assembled TEL-ON. Support **10** (0.18-0.20 mequiv/g of OH groups), prepared by reacting a carboxyl-functionalized Tentagel resin (0.28 mequiv/g of COOH groups) with 1,3-dihydroxypropane in the presence of MSNT and NMI, was used in two synthetic pathways leading to the TEL-ONs **3** and **4**. Support **10**, by reaction with the phosphoramidite linker **6b** yielded the asymmetrical protected bifunctional support **11** (0.17 mequiv/g by DMT test) which, after removal of DMT protection from the branch B and successive reaction with linker **6a** furnished the three-functionalized support **12**. Six coupling cycles with 3'-phosphoramidite units, performed as described for **7**, and a final capping step of the 5'-OH ON ends, gave **13**. Then the 5'-Fmoc group was removed from the branch A by piperidine treatment, and the successive coupling with **6a** yielded **14**. On **14** the second ON domain was assembled using 5'-phosphoramidite units, thus obtaining polymer bound TEL-ON **15**. The synthetic pathway to obtain **4** started with reaction of **10** with **6a** and subsequently **6b** as described above, thus obtaining the tetra-branched support **16** (0.28 mequiv/g by DMT). The removal of DMT groups, followed by ON synthesis with 3'-phosphoramidite units, allowed the assembly of the

first pair of ON chains. After Fmoc deprotection, the remaining ON pair having opposite polarity was then assembled using 5'-phosphoramidite units, thus obtaining **17**. Detachment from the solid support and complete deprotection of the TEL-ONs **3** and **4** were achieved treating supports **15** and **17** with concentrated  $\text{NH}_4\text{OH}$ , as described for **1** and **2**. Analyses and purifications of crude products **1-4** were carried out by HPLC using an RP18 column. The purity of the products was checked by PAGE, and the structures were confirmed by  $^1\text{H}$  NMR and MALDI-MS data. The quadruple helices were obtained by dissolving the samples **1-4** in  $\text{Na}^+$  or  $\text{K}^+$  buffers and by the annealing procedure to ensure the correct formation of quadruplex structures.

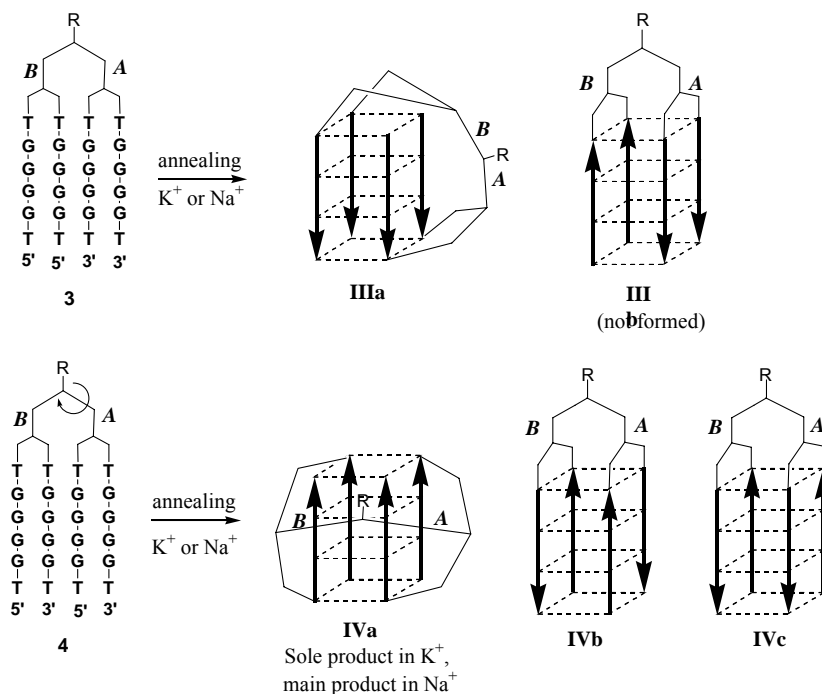
**Native Gel Electrophoresis.** The purity of the products **1-4** and the molecularity of the complexes obtained were investigated by performing a non-denaturing polyacrylamide gel electrophoresis (figure 3.9) in comparison to the tetramolecular quadruplex  $[\text{d}(\text{TG4T})]_4$ . The samples (300  $\mu\text{M}$  quadruplex) were annealed in  $\text{Na}^+$  buffer, and the electrophoretic analysis was carried out using a 20% polyacrylamide gel in TBE buffer. TEL-quadruplexes **I-IV** and  $[\text{d}(\text{TG4T})]_4$  migrated as single bands demonstrating quite similar mobilities. This result, besides indicating the purities of **I-IV**, strongly suggests that they form monomeric species at this concentration.



**Figure 3.9:** I-IV non denaturing gel analysis.

**CD and CD Thermal Analysis on Quadruplexes I-IV.** To investigate whether TEL-ONs **2-4** adopt quadruplex structures and, if such is the case, determine their nature and stability, CD studies, including CD thermal denaturation experiments, were carried out and compared with those already reported for **1** and their tetramolecular counterpart  $[\text{d}(\text{TG4T})]_4$ . The ON samples were analyzed in  $\text{Na}^+$  or  $\text{K}^+$  buffers after the annealing procedure. CD spectra (figure 3.7) were performed at 25 and 90  $^{\circ}\text{C}$  at a quadruplex concentration of  $1.0 \cdot 10^{-5}$  and  $5.0 \cdot 10^{-5}\text{M}$ . TEL-ONs **1**, **2** and  $[\text{d}(\text{TG4T})]_4$  showed almost identical CD profiles in  $\text{K}^+$  buffer at 25  $^{\circ}\text{C}$ , characterized by positive and negative bands centered around 264 and 244 nm, respectively, which are indicative of a parallel stranded G-quadruplex structure. Upon the basis of these findings, it is reasonable to hypothesize that the **2**, bearing the tetra-end-linker at the 5'-end of quadruplex, adopts the parallel G-quadruplex structure **II**. The data also suggest that **3** and **4** adopt parallel G-quadruplex structures which, considering the mutual antiparallel orientation of the ON pairs, can be explained only by hypothesizing that a considerable folding of the linker around the whole G-quadruplex takes place. The parallel structures adopted by **3** and **4** were substantiated by molecular modeling studies (see below) that clearly indicate that these

conformations are obtainable and stable as a result of the linker arrangement as shown in figure 3.10 and schematically depicted in **IIIa** and **IVa** (Scheme 3.3).



**Scheme 3.3:** Quadruplexes III and IV from bunch-ONs 3 and 4 after annealing procedure in K<sup>+</sup> or Na<sup>+</sup> buffer.

CD spectra of **1-3** and [d(TG4T)]4 in Na<sup>+</sup> buffer at 25 °C match those observed in K<sup>+</sup> showing positive and negative bands at 262 and 244 nm, respectively, whereas in the CD profile of **4** in Na<sup>+</sup> buffer a slight shift of the positive and negative bands at 259 and 239 nm, respectively, takes place. Furthermore, in the latter the presence of a new lower positive band centered at 289 nm is observable. These data indicate that even in Na<sup>+</sup> buffer, possessing a lower stabilizing effect on G-quadruplex structures when compared to K<sup>+</sup> buffer, the TEL-ON **1-3** are structured in parallel quadruplexes. For **4** in Na<sup>+</sup> buffer the coexistence of two or more conformational species could be a hypothesis, among which the parallel (main) **IVa** and the antiparallel (minor) **IVb,c** structures are present, considering that most antiparallel G-quadruplexes are characterized by a positive band centered at 290-295 nm and a negative band centered around 265 nm. CD spectra of **1-4** and [d(TG4T)]4 at 90 °C for both buffers are reported in figure 3.7. The data in K<sup>+</sup> buffer confirm that the parallel G-quadruplex structures are completely retained in the complexes **I** and **II** while a detectable destructure can be observed for complexes **III** and **IV**. CD spectra in Na<sup>+</sup> at 90 °C suggest that complexes **I-III** retain, to some extent the characteristic parallel G-quadruplex profile, while complex **IV** and [d(TG4T)]4 are completely unstructured to random coils. CD thermal denaturation-renaturation experiments were performed monitoring the CD value (mdeg) at 265 nm in the range 25-90 °C with 1 °C/min temperature scan rate. CD melting curves in K<sup>+</sup> buffer for quadruplexes **I**, **II** and d[(TG4T)]4 underwent irrelevant or low CD<sub>265</sub> variations up to 90 °C while a relatively high CD<sub>265</sub> variation was shown by the complexes **III** and **IV**. The melting profiles of the complexes **I-IV** and [d(TG4T)]4 in Na<sup>+</sup> showed higher CD<sub>265</sub> variation with

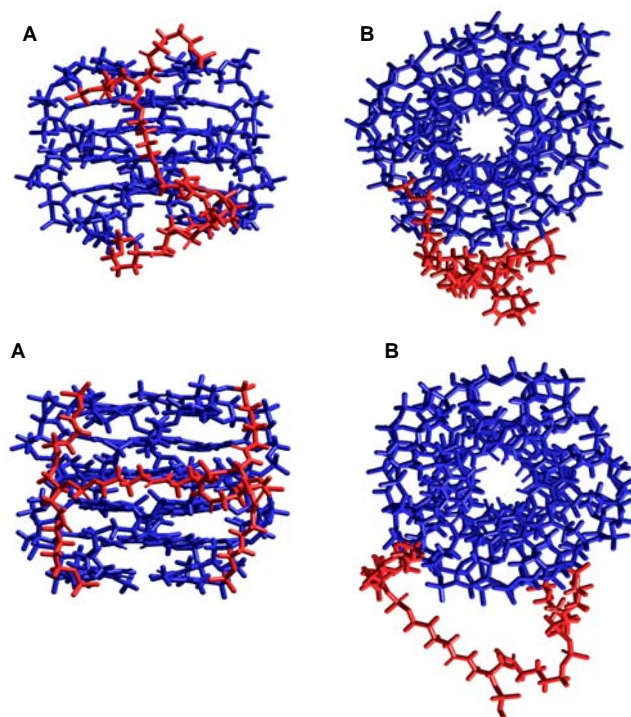
respect to the corresponding values registered in  $K^+$ ; however, only for the  $[d(TG4T)]_4$  complex was a typical and derivatizable sigmoidal curve ( $T_m=59^\circ\text{C}$ ) registered. These data indicate that the melting process for the complexes **I-III** in  $Na^+$  is not complete at  $90^\circ\text{C}$ , while a complete melting of **IV** can be observed in  $Na^+$  at this temperature. However, considering the shape of the apparent sigmoidal curve and the possible polymorphism of **IV**, we could not determine its melting temperature in both buffers. We also registered the  $CD_{265}$  thermal renaturation curves for all complexes in the range  $90-25^\circ\text{C}$ . For the tetramolecular  $[d(TG4T)]_4$  complex, as expected, a very slow renaturation process and a high hysteresis phenomenon was indicated by the profile of the renaturation curve at  $1^\circ\text{C}/\text{min}$  scan rate. The renaturation curves of **I** and **II** showed relatively low hysteresis at the same temperature scan rate. The fast reformation of the complexes **I** and **II** was also confirmed by the CD spectrum profiles registered at the end of the renaturation processes (data not shown). In the case of complexes **III** and **IV** the renaturation curves showed high hysteresis associated with low restructuring levels.

#### **$^1\text{H}$ NMR Studies on Quadruplexes I-IV and $[d(TG4T)]_4$ .**

$^1\text{H}$  NMR studies on quadruplexes **I-IV** were performed, after annealing, in the same buffers used for CD measurements. The spectrum of **I** and **II**, recorded in  $Na^+$  buffer at  $25^\circ\text{C}$  (Figure 3.8A), showed the presence of four singlets in the range 10.9-11.8ppm, two of them partially overlapped, attributable to exchange protected imino protons involved in Hoogsteen N(1)/O(6) hydrogen bonds of G-quartets. These signals were almost superimposable with those observed in the spectrum of the  $[d(TG4T)]_4$  complex already characterized by NMR data. The  $^1\text{H}$  NMR spectra recorded in  $Na^+$  buffer at  $85^\circ\text{C}$  (Figure 3.8B) indicated that **I** is completely stable at least up to  $85^\circ\text{C}$ . **II** is partially preserved (only two imino proton signals are still detectable at this temperature), whereas the tetramolecular complex  $[d(TG4T)]_4$  is completely unstructured as indicated by the disappearance of all the imino proton signals. Upon the basis of these data, it appears that quadruplex **II** in  $Na^+$  buffer is more stable than the tetramolecular counterpart, yet it is less stable than quadruplex **I**. The  $^1\text{H}$ -NMR spectra of **I**, **II**, and  $[d(TG4T)]_4$  recorded in  $K^+$  buffer displayed quite similar signal patterns; however, it was not possible to observe any quenching of the imino proton signals up to  $85^\circ\text{C}$  (Figure 3.8D). This observation is in accordance with reported data about the capability of  $K^+$  cations to form more stable quadruplex structures in comparison with those observed in  $Na^+$  buffer. As expected,  $^1\text{H}$ -NMR spectra of **III** indicated that this compound is more structured in  $K^+$  than in  $Na^+$  buffer. In  $K^+$  buffer it was possible to detect the four imino proton signals up to  $85^\circ\text{C}$  (Figure 3.8C and 3.8D), whereas in  $Na^+$  the spectrum showed large imino proton signals at  $25^\circ\text{C}$  (Figure 3.8A) which almost disappeared at  $85^\circ\text{C}$  (Figure 3.8B). The  $^1\text{H}$ -NMR spectra of **IV** at  $25^\circ\text{C}$  showed very large and overlapped imino proton signals both in  $Na^+$  (Figure 3.8A) and  $K^+$  buffers (Figure 3.8C). The increase of temperature for  $^1\text{H}$ -NMR spectra of **IV** led to the sharpening of all imino proton signals both in  $Na^+$  and  $K^+$  buffers. In  $K^+$  buffer at least three imino proton signals are clearly detectable even at  $85^\circ\text{C}$  (Figure 3.8D), whereas in  $Na^+$  buffer all imino proton signals are completely lost since  $65^\circ\text{C}$  (data not shown). These data, which are consistent with UV and CD experiments, indicate that both complexes **III** and **IV** result in less stable quadruplex structures than **I** and **II**. Furthermore, **III** results more stable than **IV** especially in  $K^+$  buffer.



**Molecular Modeling.** To explore the capability of TEL-ONs **3** and **4** of adopting the parallel quadruplex structure **IIIa** and **IVa**, respectively, a modeling study was performed. The minimized structures obtained are shown in figure 3.10.



**Figure 3.10:** Molecular models of quadruplex structures **IIIa** and **IVa**; **A**: front view, **B** top view. The tetrabranch linker is in red.

Inspection of these structures reveals that in both **IIIa** and **IVa**, the tetra-end-linker is able to span the distance between the 5'- and 3'-ends of the strands assembled in parallel quadruplex structure. It can be noted that the orientation of the linker with respect to the helical axis is different in **IIIa** and **IVa** quadruplexes. For both the complexes the inter-atomic distances between the linker atoms and the quadruplex groove hydrogen bonding sites suggest the possibility of favorable non-bonded interactions. Although the integrity of the G-tetrads is maintained after energy minimization for both complexes, a detailed analysis of the H-bonds scheme inside the G-tetrads reveals some differences between the quadruplexes. In particular, in **IIIa**, each G-tetrad shows the same hydrogen bond scheme with an optimum distance and angle between donor and acceptor atoms, resulting in strong hydrogen bonds. Indeed in **IVa** there is a change of hydrogen-bond distances and angles between adjacent quartets, and, on average, two hydrogen bonds for each G-tetrad are lost. Furthermore, the degree of coplanarity between the G residues in each G-tetrad is greater in **IIIa** than in **IVa**. It is important to point out that the perturbation of each G-tetrad directly affects the stacking energy of the neighboring G-tetrads. In summary, the TEL-ON **3** can form a parallel quadruplex structure **IIIa** without causing any distortion in the G-tetrad region, whereas in the TEL-ON **4**, which forms the quadruplex **IVa**, a major readjustment of the G-tetrads occurs.

These findings appear consistent with the experimental CD and NMR data revealing the formation of a more stable quadruplex **III** in comparison to **IV**.

The results described in this paper confirm the potential of TEL-ONs as a new class of ON analogues prone to adopt stable intramolecular quadruplex structures. In each case the  $T_m$  values are higher in potassium- than in sodium-containing buffer. It is noteworthy that, irrespective of the ends by which the four strands are attached to the tetra-end-linker, the CD spectra of the obtained complexes are typical of parallel quadruplex structures, displaying positive maxima centered at 264 nm, in contrast to maxima at 295 nm observed for antiparallel structures. The only exception concerns the CD spectrum of **4** in  $\text{Na}^+$  buffer, in which a positive shoulder at 295 nm is present, indicating that a mixture of parallel and antiparallel forms could exist (quadruplex **IVa-c**). The whole of data, including molecular modelling calculations, suggest that (i) the tetra-end-linker used in the present study is long and flexible enough to allow the organization of the ON strands into a parallel quadruplex structure and (ii) that, at least in the presence of potassium ions, this form is favored over the antiparallel one, even if it is at the cost of a major readjustment of G-tetrads, which is the case of **IVa**. To confirm these conclusions and in order to obtain TEL-quadruplexes with suitable sequences and predetermined strand orientations useful as aptamers, decoy and molecular probes, TEL-ONs characterized by shorter branched spacers have been performed., in which shortness branches could be able into force the predetermination of quadruplex strands polarity.

## EXPERIMENTAL SESSION

### General Methods

Chemicals and solvents were purchased from Fluka-Sigma-Aldrich. Reagents and phosphoramidites for DNA syntheses were purchased from Glen Research. ON syntheses were performed on a PerSeptive Biosystems Expedite. HPLC analyses and purifications were performed by a JASCO PU2089 pumps equipped with UV detector 2075 Plus using a Merck Hibar (5  $\mu\text{m}$ , 250-10) column. CPG support **5** and linkers **6a,b** were purchased from Glen Research. Tentagel carboxy resin, used as starting material for preparation of solid support **10**, was purchased from Novabiochem. Preparation of **10** was carried out in a short column (10 cm length, 1 cm i.d.) equipped with a sintered glass filter, a stopcock and a cap. The ONs were assembled by a PerSeptive Biosystems Expedite DNA synthesizer using phosphoramidite chemistry. UV spectra were run with a Jasco V 530 spectrophotometer. CD spectra and thermal denaturation experiments were obtained with a Jasco 715 circular dichroism spectrophotometer equipped with a JASCO 505T temperature controller unit. NMR spectra were recorded on a Varian Unity Inova 500 MHz spectrometer. MALDI-TOF mass spectrometric analyses were performed on a PerSeptive Biosystems voyager-De Pro MALDI mass spectrometer using a picolinic/3-hydroxypicolinic acids mixture as the matrix. The ON concentration was determined spectrophotometrically at 260 nm and 90 °C, using the molar extinction coefficient  $57800 \text{ cm}^{-1} \text{ M}^{-1}$  calculated for the unstacked oligonucleotide by the nearest neighbor mode.

### General Procedures

**Synthesis and Purification of TEL-ONs 1 and 2.** 50 mg of support **5** (0.048 mequiv/g) was used for each synthesis in the automated DNA synthesizer following standard phosphoramidite chemistry, using 45 mg/mL of solution of phosphoramidite **6a** (two coupling cycles) and 3'- or 5'-phosphoramidite nucleotide building block (six cycles) in anhydrous CH<sub>3</sub>CN, thus obtaining the polymer bound ON **8** and **9**, respectively. The coupling efficiency with nucleotide units was consistently higher than 98% (by DMT spectrophotometric measurements). The solid support **8** (or **9**) was then treated with 25% aq ammonia solution for 7 h at 55 °C. The filtered solution and washings were concentrated under reduced pressure and purified by HPLC on a RP18 column eluted with a linear gradient CH<sub>3</sub>CN in TEAB buffer (pH 7.0, from 0 to 100% in 60 min, flow 1.0 mL/min), retention time 26.1 and 27.0 min for **1** and **2**, respectively. The collected peaks furnished pure **1** and **2** (72 and 77 OD<sub>260</sub> units, respectively) which, after lyophilization, were characterized by spectroscopic data. TEL-ON **2**: MALDI TOF-MS (negative mode): calculated mass (8763.7); found: 8762.6.

**Preparation of Support 10.** 500 mg of Tentagel carboxyl resin (0.28 mequiv/g of COOH) suspended in CH<sub>2</sub>Cl<sub>2</sub> were left dihydroxypropane (0.56 mmol, 0.36 mL), N-methylimidazole (NMI, 0.56 mmol, 0.40 mL) and mesitilensulphonylnitrotriazole (MSNT 0.056 mmol, 1.48 g) and the mixture was shaken for 36 h at room temperature. The obtained hydroxyl support **10** was exhaustively washed with CH<sub>2</sub>Cl<sub>2</sub> and then dried under reduced pressure. The amount of hydroxyl groups, determined by DMT measurements after the successive coupling step with 3'-phosphoramidite nucleotide units, resulted to be in the range 0.18-0.20 mequiv/g.

**Synthesis and Purification of TEL-ONs 3 and 4.** 50 mg of support **10** (0.18 mequiv/g) reacted, in the automated DNA synthesizer, with phosphoramidite **6b** (45 mg/mL in CH<sub>3</sub>CN) following standard phosphoramidite chemistry at 15-μmol scale, yielded support **11** (0.17 mequiv/g of DMT groups). After removal of the DMT group by DCA/standard treatment, the second coupling cycle was performed in the same manner using phosphoramidite **6a**, thus obtaining support **12** (0.30 mequiv/g of DMT groups). **12** was then subjected to six coupling cycles using 3'-phosphoramidite nucleotide building blocks (45 mg/mL in CH<sub>3</sub>CN, 15-μmol scale), followed by final DMT removal and 5'-OH capping with Ac<sub>2</sub>O, to yield ON-functionalized support **13**. The removal of Fmoc groups was achieved by treatment with piperidine/DMF solution (2:8 v/v, 30 min RT). The resulting support was then reacted with phosphoramidite **6a**, thus obtaining **14** (0.27 mequiv/g of DMT groups). On branches A of **14** the second ON domain was assembled, performing six coupling cycles with the 5'-phosphoroamidite nucleotide building block as described for support **12**, yielding the polymer bound ON **15**. For the synthesis of **4**, support **10** (50 mg) was functionalized with **6a** and then **6b**, as described before, yielding the tetrabranch support **16** (0.28 mequiv/g of DMT groups). After removal of the DMT protecting groups, the first two ON chains were assembled, using 3'-phosphoramidite nucleotide building blocks. After capping of the terminal 5'-OH functions by Ac<sub>2</sub>O treatment, the Fmoc groups were removed as previously described, and the successive two ON chains were assembled using 5'-phosphoroamidite nucleotide building blocks, thus obtaining the polymer bound ON **17**. TEL-ONs were detached from the supports **15** and **17**, deprotected, and purified as described for **1** and **2**, thus obtaining **3** and **4** (retention time 25.8 and 25.2 min, respectively). After lyophilization, the final pure products **3** and **4** (68 and 62 OD<sub>260</sub> units, respectively) were characterized by spectroscopic data. TEL-ON **3**: MALDI TOF-MS (negative mode): calculated mass (8821.7); found: 8820.5. TEL-ON **4**: MALDI TOF-MS (negative mode): calculated mass (8821.7); found: 8820.2

**Preparation of Quadruple Helices (Annealing).** Quadruplexes **I-IV** were formed by dissolving the TEL-ONs **1-4**, in the appropriate buffer and annealed by heating to 90 °C for 20 min followed by slow cooling to room temperature. The solutions were equilibrated at 10 °C for 24 h before performing the experiments. The buffers used were 10 mM Na<sub>2</sub>HPO<sub>4</sub>, 70 mM NaCl, 0.2 mM EDTA (Na<sup>+</sup> buffer) or 10 mM K<sub>2</sub>HPO<sub>4</sub> 70 mM KCl, and 0.2 mM EDTA (K<sup>+</sup> buffer).

**Native Gel Electrophoreses.** Native gel electrophoreses were run on 20% nondenaturing polyacrylamide gels in 0.6x TBE buffer, pH 7.0, with 50 mM NaCl. ON solutions, with a final concentration of 300 μM in quadruplex, were prepared in 50 mM Tris-borate pH 7.0, 80 mM NaCl, 1 mM spermine, and 3 mM MgCl<sub>2</sub> and annealed as above-described. At 2 μL of the ODN solution (300 μM in quadruplex) was then added 1 μL of 30% glycerol and the mixture charged on gel. The gels were run at room temperature at constant voltage (100 V) for 2.5 h. The bands were visualized by UV shadowing.

**CD Experiments.** CD spectra of the quadruplexes were registered on a Jasco 715 circular dichroism spectrophotometer in a 0.1 cm path length cuvette at 25 and 90 °C at 1.0 x10<sup>-5</sup> and 5.0 x10<sup>-5</sup>M quadruplex concentrations. The wavelength was varied from 220 to 320 nm at 5 nm min<sup>-1</sup>. The spectra were recorded with a response of 16 s, at 2.0 nm bandwidth and normalized by subtraction of the background scan with buffer. CD thermal denaturation-renaturation experiments were performed monitoring the CD value (mdeg) at 264 nm in the range 25-90 °C with 1 °C/min heating rate.

**<sup>1</sup>H NMR Experiments.** <sup>1</sup>H NMR data were collected on a Varian Unity INOVA 500 spectrometer equipped with a broadband inverse probe with z-field gradient and processed using Varian VNMR software package. 1D-NMR spectra were acquired as 16,384 data points with a recycle delay of 1.0 s at temperatures in the range 25-85 °C. Data sets were zero filled to 32,768 points prior to Fourier transformation and apodized with a shifted sine bell squared window function. Pulsed-field gradient WATERGATE sequence was used for H<sub>2</sub>O suppression. All NMR samples **I-IV** were prepared in H<sub>2</sub>O/

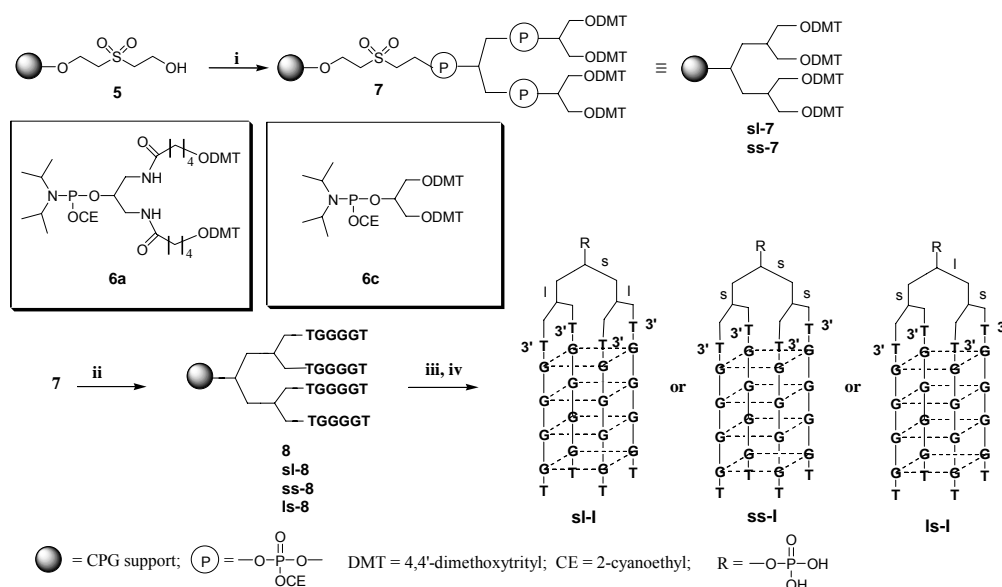
D<sub>2</sub>O (9:1, v/v) at 0.5 mM of quadruplex concentration with a final salt concentration of 100 mM KCl and 10 mM K<sub>2</sub>HPO<sub>4</sub> or 100 mM NaCl and 10 mM Na<sub>2</sub>HPO<sub>4</sub> for K<sup>+</sup> or Na<sup>+</sup> buffer, respectively.

**Molecular Modeling.** The conformational features of the quadruplexes **III** and **IV** have been explored by means of molecular modeling study. All the calculations were performed on a personal computer running the HyperChem 7.5 suite of programs. The AMBER force field based on AMBER 99 parameter set was used. The initial coordinates for the starting model of [d(TGGGGT)]<sub>4</sub> quadruplex were taken from the NMR solution structure of the [d(TTGGGGT)]<sub>4</sub> quadruplex (Protein Data Bank entry number 139D), choosing randomly one of the four available structures. The initial [d(TGGGGT)]<sub>4</sub> quadruplex models were built by deletion of the 5'-end thymidine residue in each of the four TTGGGGT strands. The complete structures of **III** and **IV** were then built using the HyperChem 7.5 building tool. Partial charges for each of the tetra-end-linker atoms were assigned using the Gasteiger-Marsili algorithm implemented in the QSAR module of HyperChem suite. The resulting coordinates of tetra-end-linker's atoms were energy minimized in a vacuum, keeping all DNA coordinates frozen (500 cycles of the steepest descent method). Three sodium ions were manually positioned in the central channel, equidistant from adjacent G-tetrads to allow octahedral coordination with G carbonyl oxygen O6 atoms. The two complexes were neutralized by addition of further 26 sodium ions placed in the most negative positions of the systems. Each quadruplex was then solvated by a periodic TIP3P water box (5833 water molecules), which extended to a distance

of 10 Å from each atom. The minimal distance between water and system atoms was set to 2.3 Å. The whole systems were initially subjected to energy minimization followed by 50 ps of molecular dynamics at 300 K, and then the systems were slowly cooled at 10°K. The obtained structures were further minimized using 1000 cycles of the steepest descent method followed by conjugate gradient method until convergence to a rms gradient of 0.1 kcal mol<sup>-1</sup> Å<sup>-1</sup>.

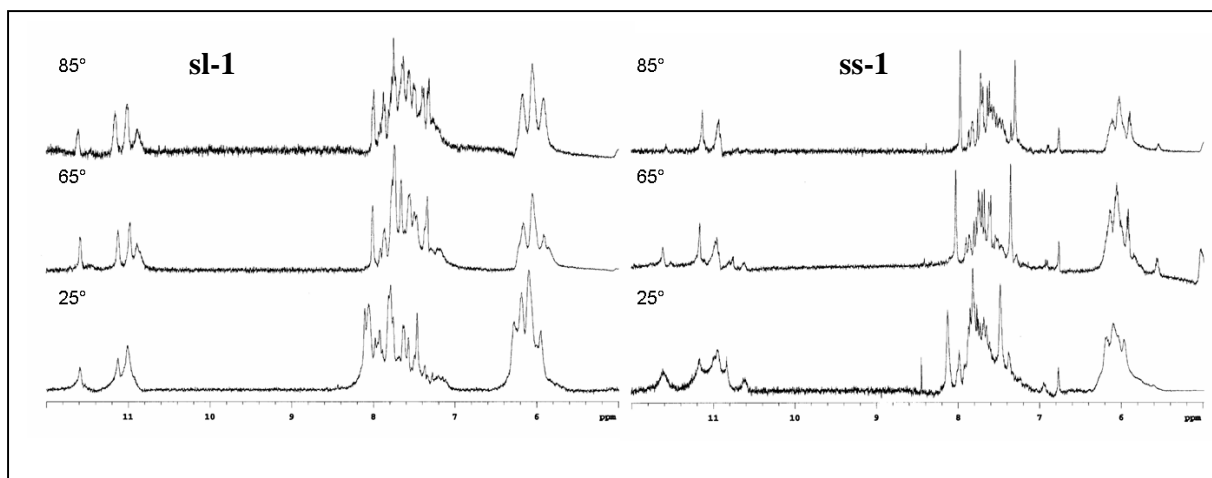
### 3.9. Studies on the length of tetra-end-linked linker on the orientation of the ON strands aimed at obtaining stable quadruplex structures

In the attempt to obtain monomolecular quadruplexes having predetermined strands polarity, the synthesis of two new tetra-branched linker moieties characterized by the shortening of the primary or both primary and secondary TEL branches (**sl-TEL** or **ss-TEL**, Scheme 3.4), was performed. To test if the **sl-TEL** and **ss-TEL** moieties were still able to allow the formation of stable quadruplex structures, synthesis of **sl-1** and **ss-1** via the above reported solid-phase procedure using the CPG support and the branching bi-functional linkers **6a** and **6c** were performed. Is-TEL-d(TGGGGT)<sub>4</sub> (**Is-1**) was not synthesized because its primary branch is long enough to span the distance between the quadruplex edges, thus allowing the loss of strands polarity predetermination.



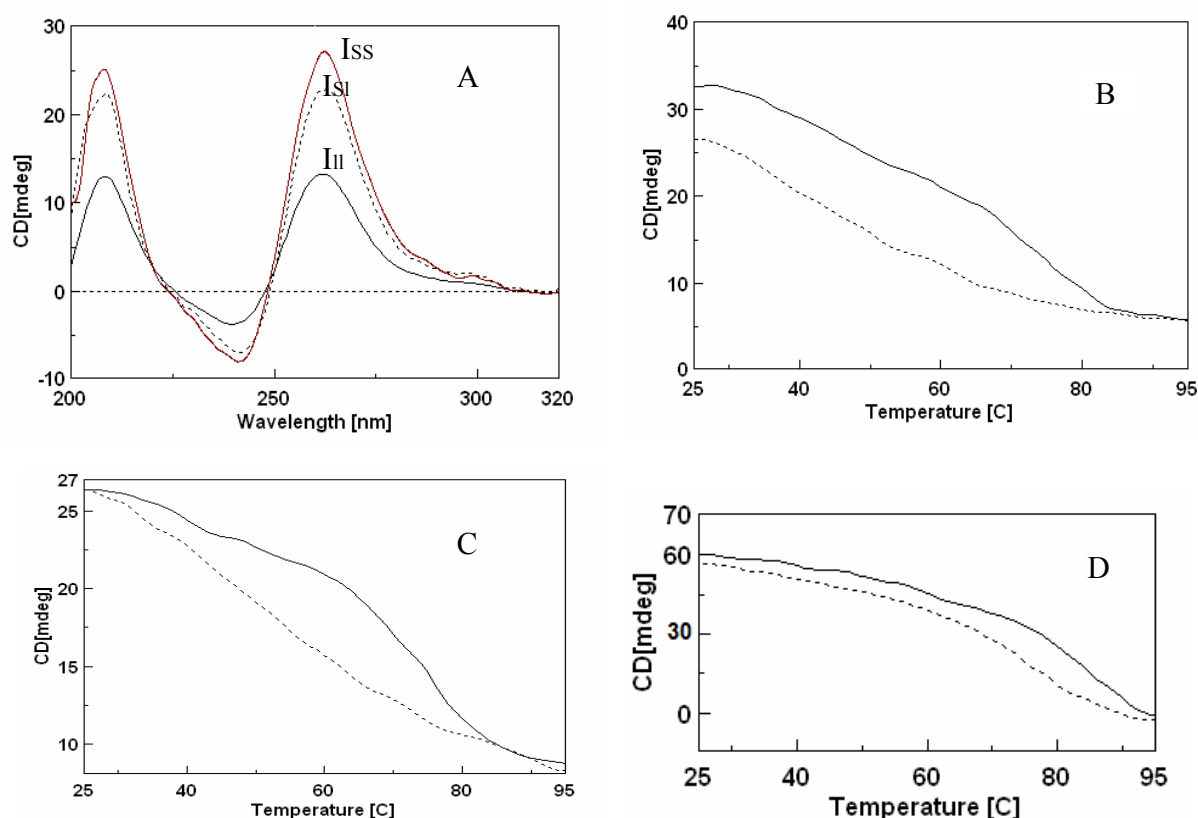
**Scheme 3.4:** i: two couplings with **6a**, **6c**; ii: ODN synthesis with 3' or 5'-phosphoramidites; iii: detachment and deprotection with NH<sub>4</sub>OH conc. 32% (7 h, 55 °C); iv: HPLC purification and annealing.

**<sup>1</sup>H-NMR and CD studies.** In order to assess if **sl-1** and **ss-1** could still fold into the quadruplex structures **sl-I** and **ss-I**, respectively, and if such were the case, to estimate their stabilities, we performed <sup>1</sup>H-NMR (at 25, 65, and 85°C using pulse field gradient WATERGATE sequences for H<sub>2</sub>O suppression) and CD studies. Comparison of <sup>1</sup>H-NMR stacked spectra of **sl-I** and **ss-I** (figure 3.11) showed that both adopted parallel quadruplex structures in presence of K<sup>+</sup> ions.



**Figure 3.11:**  $^1\text{H}$  NMR (500 MHz) spectra of **sl-I** and **ss-I** and  $[\text{d}(\text{TG4T})]_4$  in  $\text{K}^+$  buffers. The spectra were registered in  $\text{H}_2\text{O}/\text{D}_2\text{O}$ , (9:1, v/v) solution at 0.5 mM of quadruplex at 25, 65 and 85 °C. The aromatic and imino proton regions of the spectra are reported.

However, it appeared that the presence of the ss-TEL moiety lead to the lowering of the thermal stability of **ss-I**, and to the partial loss of the magnetic equivalence between the imino protons in each G-tetrad. On the other hand **sl-I**, characterized by the presence of the sl-TEL moiety, appeared to be very stable up to 85°C. Furthermore, in the case of **sl-I** no loss of imino protons magnetic equivalence was observed. Finally, the CD spectra and the CD melting curves were well in agreement with the NMR data (figure 3.12). Upon the basis of these findings, the sl-TEL seems to be very promising for the achieving of monomolecular multistranded quadruplex structures characterized by both high thermal stability and predeterminable strands orientation. CD studies suggest that the shortening of the branches could enhance the folding of the quadruplex: i) **ss-I** appears more folded (higher mdeg value at 264nm) than **sl-I**, followed by **I** (overlayer spectra, samples prepared in the same concentration and ionic strength conditions); ii) the melting process for all complexes in study is not complete at 90°C; iii) thermal renaturation curves have been registered for all the complexes in the range 20-95°C. Data indicate that **sl-I** such as **I** (figure 3.12C and 3.12B, respectively) showed relatively low hysteresis at 1°C/min scan rate.



**Figure 3.12:** A) Overlay CD spectra of I (I<sub>II</sub>) sl-I and ss-I at 25°C. CD melting (solid line) and refolding (dotted line) for each one I<sub>II</sub> (B), sl-I (C) and ss-I (D) in the range 25-90°C at 1°C/min heating rate.

The fast reformation of the complex **ss-I** (figure 3.12D) and the absence of hysteresis in its case suggest that **ss-I** could be a good candidate to realize studies on nanomotors. Furthermore, CD studies on the refolding rate of **I**, **sl-I** and **ss-I**, have been made (data not shown). **ss-I** refolding is complete in just 15 min (the shortest time to register melting and refolding curves in 25-95°C range by CD instrument), while **I** and **sl-I** are not completely refolded in the same conditions. The syntheses of the antiparallel sl-TEL-d(TGGGGT)<sub>4</sub> (**sl-III** and **sl-IV**) and also ss-TEL-d(TGGGGT)<sub>4</sub> (**ss-III** and **ss-IV**) have been performed. These structural features may be able to hinder the folding of the linker around the whole quadruplex structure thus forcing the predetermination of quadruplex strands polarity upon the basis of the relative orientation of the ON strands into the TEL-ON molecules. CD and NMR data on these new synthesized TEL-ON analogues are currently acquiring.

## EXPERIMENTAL SESSION

### General Procedures

The **sl-I** and **ss-I**; **sl-III** and **ss-III**; **sl-IV** and **ss-IV** syntheses (scheme 3.5) were performed using the same protocols employed for the corresponding analogues above reported (**I,III,IV**). The only difference is in TEL-supports synthesis. CD and NMR studies were performed using the same procedures for the **1-4** analogues above reported.





**Synthesis and Purification of sl-3 and ss-3 TEL-ONs.** For the syntheses of **sl-3** and **ss-3**, 50 mg of support **10** (0.18 mequiv/g) in each case were reacted, in the automated DNA synthesizer, with phosphoramidite **6d** (45 mg/mL in CH<sub>3</sub>CN) following standard phosphoramidite chemistry at 15- $\mu$ mol scale, yielded support **s-11** (0.17 mequiv/g of DMT groups). After removal of the DMT group by DCA/standard treatment, the second coupling cycle was performed in the same manner using phosphoramidite **6a** or **6c**, respectively to obtain support **sl-12** or **ss-12**, respectively (0.30 mequiv/g of DMT groups). These latters were then subjected to six coupling cycles using 3'-phosphoramidite nucleotide building blocks (45 mg/mL in CH<sub>3</sub>CN, 15- $\mu$ mol scale), followed by final DMT removal and 5'-OH capping with Ac<sub>2</sub>O, to yield ON-functionalized supports **sl-13** from **sl-12** and **ss-13** from **ss-12**. The removal of Fmoc groups was achieved by treatment with piperidine/DMF solution (2:8 v/v, 30 min RT). The resulting support was then reacted with phosphoramidite **6a** or **6c** respectively, thus obtaining **sl-14** and **ss-14** (0.27 mequiv/g of DMT groups). On these latters, the second ON domain was assembled, performing six coupling cycles with the 5'-phosphoroamidite nucleotide building block as described for support **12**, yielding the polymer bound **sl-ON 15** and **ss-ON 15**.

**Synthesis and Purification of sl-4 and ss-4 TEL-ONs.** Support **10** (50 mg) was functionalized with **6c** and then **6b**, and/or **6c** followed by **6d**, yielding the tetra-branched supports **sl-16** and/or **ss-16**, respectively (0.28 mequiv/g of DMT groups). After removal of the DMT protecting groups, the first two ON chains were assembled, using 3'-phosphoramidite nucleotide building blocks. After capping of the terminal 5'-OH functions by Ac<sub>2</sub>O treatment, the Fmoc groups were removed as previously described, and the successive two ON chains were assembled using 5'-phosphoroamidite nucleotide building blocks, thus obtaining the polymer bound **sl-ON 17** and/or **ss-ON 17**.

TEL-ONs were detached from the supports **sl-15** and **ss-15** and **sl-17** and **ss-17**, deprotected, and purified as described for **3** and **4**, thus obtaining **sl-3** and **ss-3** and **sl-4** and **ss-4** (retention time 25.3 and 24.5 min, respectively for **sl-3** and **ss-3**; 24.7 and 24.2 min, respectively for **sl-4** and **ss-4**). After lyophilization, the final pure products **sl-3** and **ss-3** (75 and 78 OD<sub>260</sub> units, respectively) and **sl-4** and **ss-4** (65 and 80 OD<sub>260</sub> units, respectively) were characterized by spectroscopic data. **sl-TEL-ON 3**: MALDI TOF-MS (negative mode): calculated mass (8320.26); found: 8320.5. **ss-TEL-ON 3**: MALDI TOF-MS (negative mode): calculated mass (8061.7); found: 8062.6. **sl-TEL-ODN 4**: MALDI TOF-MS (negative mode): calculated mass (8320.26); found: 8319.2. **ss-TEL-ODN 4**: MALDI TOF-MS (negative mode): calculated mass (8061.7); found: 8062.6.

**Synthesis of 6c.** 0.34g (3.6mmol) of 1,2,3-propantriol dried by repeated coevaporations with anhydrous pyridine and then dissolved in 8mL of the same solvent, were reacted with 2.5g (7.2mmol) of 4,4'-dimethoxytrytilchloride at R.T. for 12h. The reaction was quenched by addition of water and the resulting mixture, taken to dryness and redissolved in benzene, was purified on a silica gel column giving 1.63 g (2.34 mmol, 65% yield) of 1,3- bis-(O-DMT)-propantriol, in mixture with the full protected derivative. The bis-tritylated compound was dissolved in dry CH<sub>2</sub>Cl<sub>2</sub> (5mL) and N,N-diisopropylethylamine was added (0.860 mL, 4.6 mmol) followed by 2-cyanoethyl-N,N-diisopropylchlorophosphoramidite (0.48 mL, 2.76 mmol). The mixture was stirred for 1h at R.T. then diluted with ethyl acetate (50mL) and washed with saturated sodium chloride solution(25mLx4).The organic phase was dried over sodium sulphate and evaporated to an oil in the presence of toluene(10mL). The product was purified by short column chromatography on Kieselghel 60H (10g) eluting with CH<sub>2</sub>Cl<sub>2</sub>/ethyl acetate/2,6-lutidine 50:50:2. The product was dried to give a white powder (90%yields).

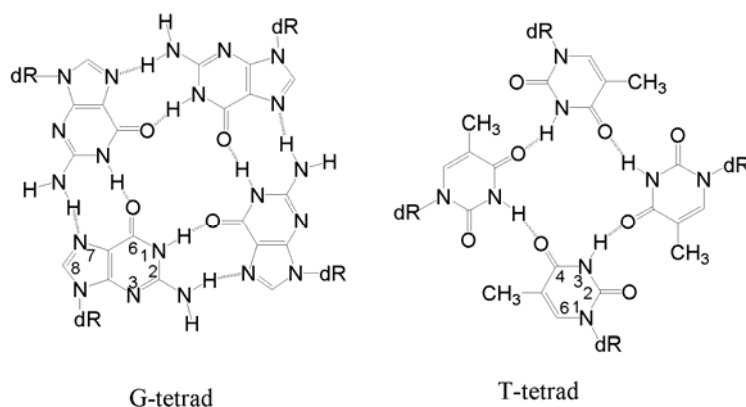
**6c**  $^1\text{H-NMR}$  (400 MHz) ppm (acetone,  $\text{d}_6$ ) 1.05 (d, 12H,  $\text{CHCH}_3$ ), 2.97 (m, 2H,  $\text{CHCH}_3$ ), 3.52 (d, 4H,  $\text{CH}_2\text{ODMT}$ ), 3.82 (s, 12H,  $\text{OCH}_3$ ), 3.87 (m, 1H,  $\text{CHOP}$ ), 4.05 (s, 2H,  $\text{CH}_2\text{CN}$ ), 6.82-7.55 (m, 26H, aromatic protons). ESI-MS calculated  $m/z$ : 882.11, ( $\text{C}_{53}\text{H}_{59}\text{N}_2\text{O}_8\text{P}$ ), found: 883.40 ( $\text{M} + \text{H}$ ) $^+$

**Synthesis of 6d.** 0.34 g (3.6 mmol) of 1,2,3-propantriol dissolved in 8 mL of pyridine, were reacted with 0.325 g (0.9 mmol) of 4,4'-dimethoxytrytilchloride at R.T. for 12 h to afford after purification the monotrityl derivative as main product (0.343 g, 0.8 mmol, 80%). This latter was protected at the remaining primary hydroxyl group by reaction with fluorenylmethoxycarbonylchloride (Fmoc-Cl 232 mg, 0.9 eq) in anhydrous  $\text{CH}_2\text{Cl}_2$ . Phosphoramidite derivative was prepared as just described for obtaining 6c.

**6d**  $^1\text{H-NMR}$  (400 MHz) ppm (acetone,  $\text{d}_6$ ): 1.05 (d, 12H,  $\text{CHCH}_3$ ), 2.97 (m, 2H,  $\text{CHCH}_3$ ), 3.52 (d, 2H,  $\text{CH}_2\text{ODMT}$ ), 3.82 (s, 6H,  $4\text{OCH}_3$ ), 4.05 (s, 2H,  $\text{CH}_2\text{CN}$ ), 4.14 (m, 1H,  $\text{CHOP}$ ), 4.31 (d, 2H,  $\text{CH}_2\text{OFmoc}$ ), 4.46 (t, 1H,  $\text{CH-fluorenyl}$ ), 4.78 (d, 2H,  $\text{CH}_2\text{-fluorenyl}$ ), 6.80-7.65 (m, 21H, aromatic protons). ESI-MS calculated  $m/z$ : 802.34, ( $\text{C}_{47}\text{H}_{51}\text{N}_2\text{O}_8\text{P}$ ), found: 803.80 ( $\text{M} + \text{H}$ ) $^+$ .

### 3.10 Studies on quadruplex complexes, containing less stable quartets such as T-tetrads

In order to extend TEL-ON approach in the study of quadruplex complexes containing less stable quartets, T-tetrads have aroused interest since methyl groups of T residues could give rise to hydrophobic contacts into G-quadruplex grooves potentially useful for improving the interaction between quadruplex-based aptamers and their target molecules. Even though the presence of a T-tetrad in quadruplex structures could be an attractive topic, particularly in the aptamer research area, it shows a main drawback due to the minor stability of a T-tetrad (only four hydrogen bonds and a smaller surface for stacking interactions in comparison with a G-tetrad, figure 3.13) that could decrease the stability of the overall complex. Since TEL-ON analogue is provided with more favourable thermodynamic properties than its tetramolecular counterpart, it is reasonable to suppose that the destabilization of a quadruplex structure due to the presence of a T-tetrad could be counterbalanced by the improvement caused by the tetra-banched spacer. In order to verify this hypothesis, the synthesis of the  $[\text{d}(\text{TG}2\text{TG}2\text{C})]_4$ -bunch (**1<sub>T</sub>**) (Scheme 3.6) has been performed, whose base sequence matches that of the ON **3<sub>T</sub>** forming the tetramolecular quadruplex structure **III<sub>T</sub>** in which a T-tetrad was observed for the first time<sup>[90]</sup>.



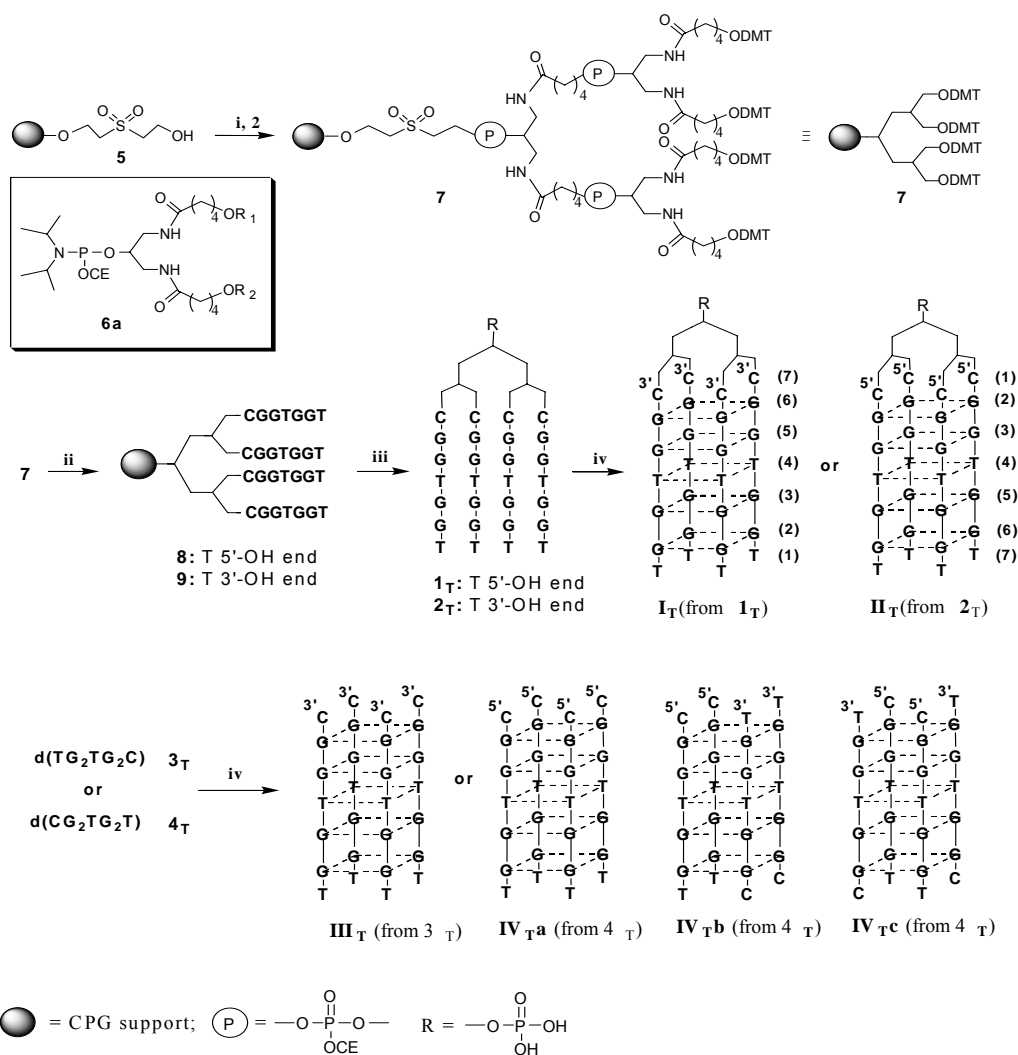
**Figure 3.13:** Hydrogen bonds involved in G and T tetrads

The resulting  $I_T$  was investigated by circular dichroism (CD), CD thermal denaturation experiments,  $^1H$ , and nuclear Overhauser effect spectroscopy (NOESY) nuclear magnetic resonance (NMR) experiments. Further, studies on the TEL-[d(CG2TG2T)]<sub>4</sub> ( $2_T$ ) and its corresponding natural ON  $4_T$ , have been carried out, in which, in respect to  $1_T$ , the C and T end bases are mutually exchanged, with the aim to explore the capability of the tetrabranched linker to stabilize other sequences potentially prone to form quadruplex structures containing a T-tetrad.

**Synthesis and Purification of Quadruplexes  $I_T$ – $IV_T$ .** The syntheses of TEL-ONs  $1_T$  and  $2_T$  (Scheme 3.6) have been performed using the above reported solid-phase procedure, which uses the hydroxyl functionalized CPG support **5** and the bifunctional linker **6a** to prepare support **7** bearing a symmetrical tetrabranched linker with four protected primary alcoholic functions, prone to ON chain assembly. Support **7** was used to synthesize the TEL-ONs  $1_T$  and  $2_T$ , each characterized by having the four ON strands linked to the spacer through the 3'- or 5'-end, respectively. The ON chains were assembled by an automatic DNA synthesizer using 3'- or 5'-phosphoramidite nucleotide building blocks, thus obtaining the polymer bound TEL-ONs **10** and **11** respectively.

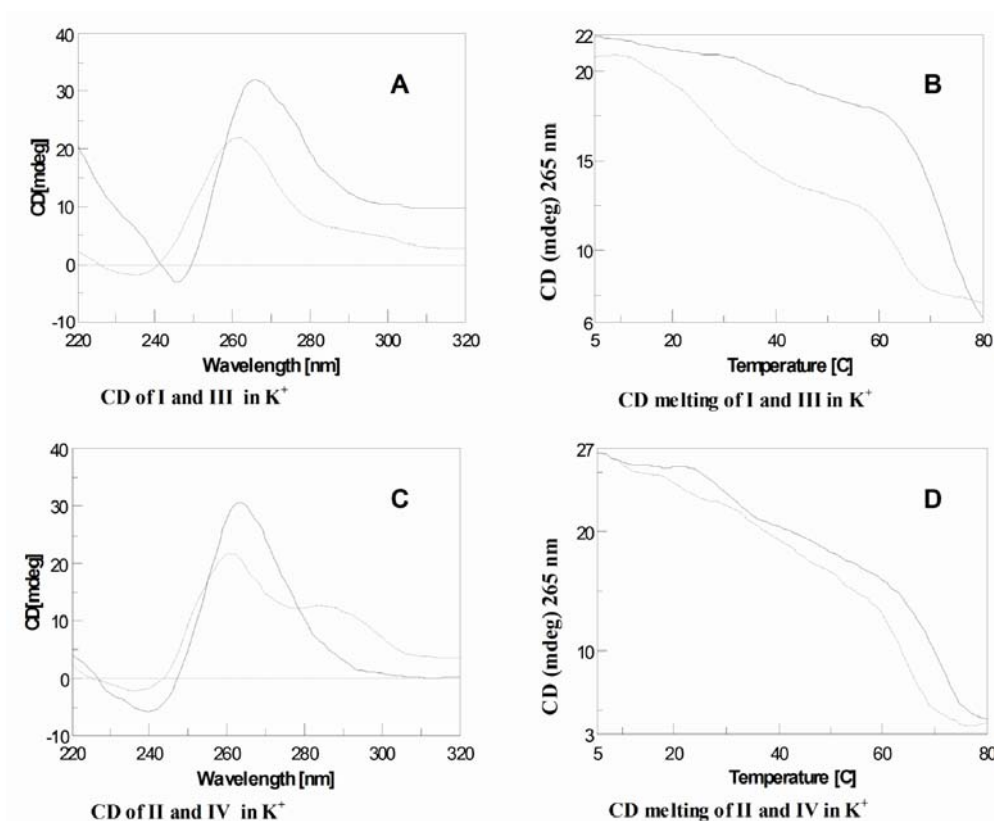
Supports **10** and **11** treated with conc.  $NH_4OH$  (7 h,  $55^\circ C$ ), afforded crude TEL-ONs  $1_T$  and  $2_T$ , which after high performance liquid chromatography (HPLC) purification and annealing procedure furnished the quadruplexes  $I_T$  and  $II_T$ , respectively. The natural ONs  $3_T$  and  $4_T$ , precursors of the tetramolecular quadruplexes  $III_T$  and  $IV_T$ , respectively, have been prepared using a standard automatic DNA synthetic procedure followed by HPLC purification. The purity of the products was checked by polyacrylamide gel electrophoresis (PAGE) and the structures confirmed by  $^1H$ -NMR and matrix-assisted laser desorption ionization–mass spectroscopy (MALDI-MS) data. The quadruple helices  $I_T$ – $IV_T$  were obtained dissolving the samples  $1_T$ – $4_T$  in  $K^+$  buffers (110 mM  $K^+$ ) and annealing to assure the correct formation of quadruplex structures.

**CD Experiments.** To investigate whether TEL-ONs  $1_T$  and  $2_T$  adopt the quadruplex structures  $I_T$  and  $II_T$ , and to determine their nature and stability in comparison with the tetramolecular complexes  $III_T$  and  $IV_T$ , CD studies, including CD thermal denaturation experiments were carried out after the annealing procedure. CD spectra (Figure 3.14) were performed at  $25^\circ C$  at a quadruplex concentration of  $2.0 \cdot 10^{-5}$  M in  $K^+$  buffer. TEL-ONs  $1_T$ ,  $2_T$ , as well as natural ON  $3_T$ , showed very similar CD profiles characterized by positive bands at 266, 264, and 262 nm, respectively, and negative bands at 246, 240, and 238 nm, respectively (figure 3.14 A and 3.14C), which are indicative of parallel stranded G-quadruplex structures ( $I_T$ ,  $II_T$ , and  $III_T$ ). However, the CD spectrum of  $4_T$  showed a different profile characterized by the presence of two positive bands at 262 and 289 nm and a negative band at 239 nm (figure 3.14C). To further investigate the unexpected CD profile showed by  $4_T$ , we registered CD spectra in the range  $5$ – $80^\circ C$  (melting process) and  $80$ – $5^\circ C$  (fast annealing process) at  $1^\circ C/min$  scan rate (figure 3.15A and 3.15B, respectively). In each process, the CD values of the positive bands at 262 and 289 nm varied in a non constant ratio. These data suggest that the two positive bands can be attributed to different complexes. Considering that the majority of publications have reported that most antiparallel G-quadruplexes are characterized by a positive band centred at 290–295 nm and a negative band centred around 265 nm, it is reasonable to hypothesize that  $4_T$  leads to a mixture of quadruplex species probably containing parallel ( $IV_{Ta}$ ) and antiparallel complexes ( $IV_{Tb}$ ,  $IV_{Tc}$ ).



**Scheme 3.6:** i: two couplings with 6a; ii: ON synthesis with 3' or 5'-phosphoramidites; iii: detachment and deprotection with  $\text{NH}_4\text{OH}$  conc. 32% (7 h,  $55^\circ\text{C}$ ); iv: HPLC purification and annealing.

CD Thermal Analysis and CD thermal denaturation experiments were performed in  $\text{K}^+$  buffer monitoring the CD value (mdeg) at 265 nm in the range  $5\text{--}80^\circ\text{C}$  with  $1^\circ\text{C}/\text{min}$  heating rate. The melting curve of  $\text{I}_T$  (figure 3.14B) did not show a well-defined change of convexity up to  $80^\circ\text{C}$ ; however, a drastic reduction of the  $\text{CD}_{265}$  value was observed in the range  $60\text{--}80^\circ\text{C}$ . The contrary is true in the melting profile of the tetramolecular complex  $\text{III}_T$  (figure 3.14D), since  $15\text{--}20^\circ\text{C}$ , a significant reduction of the  $\text{CD}_{265}$  value, and two detectable changes of convexity were observed. These data suggest that the helix–coil transition point for complex  $\text{I}_T$  is at a higher temperature and is more cooperative than that of the corresponding tetramolecular complex  $\text{III}_T$ .

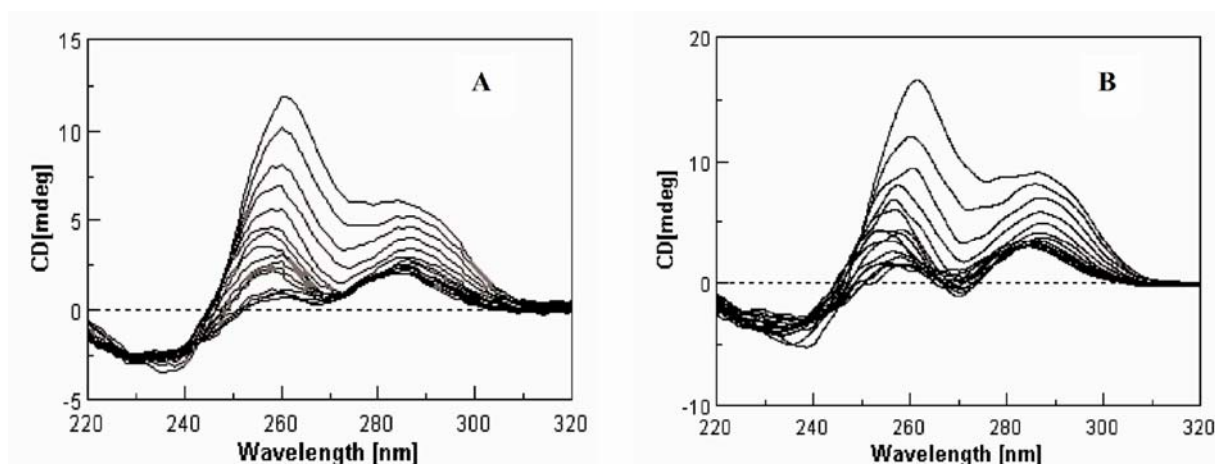


**Figure 3.14:** CD spectra at 5°C and melting profiles of quadruplexes I–II (solid line) and III–IV (dotted line) in  $K^+$  buffer.

The melting curve of  $II_T$  was shown to be very similar to that of  $IV_T$  (figure 3.14D), both being characterized by the reduction of  $CD_{265}$  values up to 55–60°C followed by a point of inflection and the complete melting at 80°C, even if a shift toward higher temperature is clearly noticeable for the CD profile of  $II_T$ . The whole of the CD melting data suggests that TEL-ONs  $I_T$  and  $II_T$  are more stable than their natural counterparts  $III_T$  and  $IV_T$ , respectively, although a quantitative determination of the relative thermal stability was not performed yet, due to the very slow kinetics of quadruplex denaturation process and considering that, in the case of  $II_T$  and  $IV_T$ , the latter could be a mixture of different complexes.

**$^1H$ -NMR Studies on Quadruplexes  $I_T$ – $IV_T$ .** Temperature-dependent one-dimensional spectra (25, 45, 65, and 75°C) and NOESY spectra (25°C) were recorded in  $H_2O/D_2O$  (9:1, v/v) in  $K^+$  buffer using pulse field gradient WATERGATE for  $H_2O$  suppression.

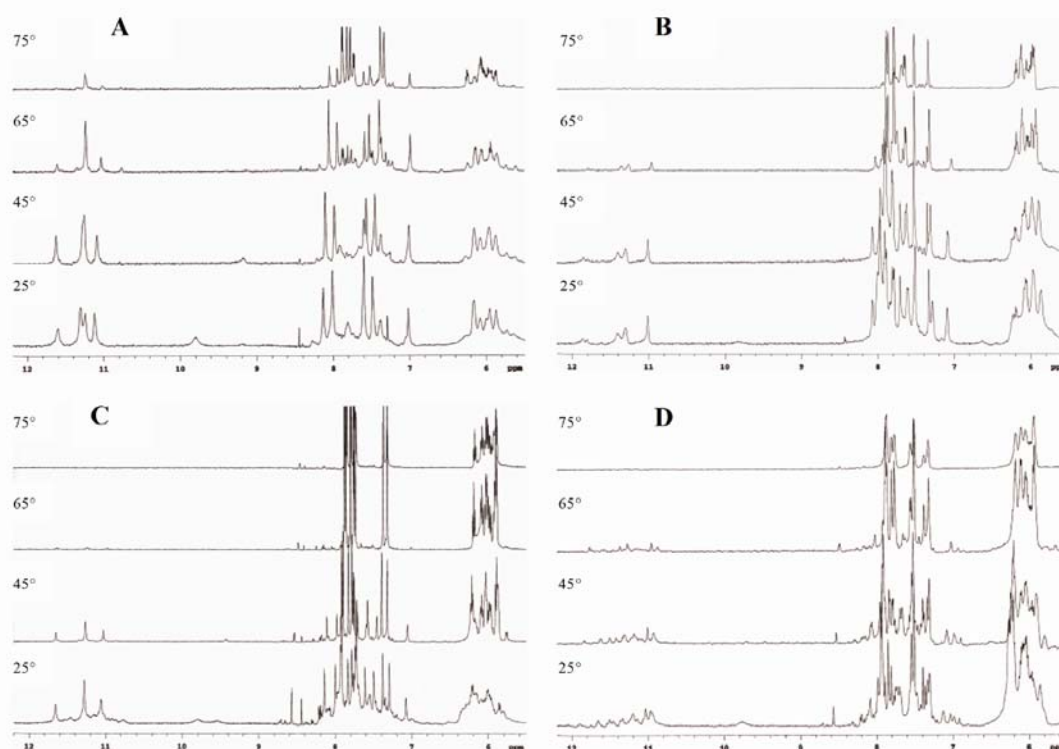
$^1H$ -NMR spectrum of complex  $I_T$  at 25°C (figure 3.16A) showed the presence of four well-resolved signals between 11.1 and 11.6 ppm attributable to exchange protected imino protons involved in Hoogsteen N(1)H/O(6) hydrogen bonds of G-quartets also on the basis of the comparison with the corresponding region of the spectrum of  $III_T$  (figure 3.16C) that forms a stable parallel stranded quadruplex already characterized by NMR data.



**Figure 3.15:** CD spectra of IV. (A) Melting process 5–80°C; (B) fast annealing process 80–5°C, at 1°C/min scan rate.

In  $I_T$ , the broad signal centred around 9.8 ppm is attributable to H-bonded GNH2 protons, while the weak signal at 9.2 ppm is attributable to T4NH imino protons involved in N(3)H/O(4) hydrogen bonds, thus confirming the formation of a T-tetrad comprised between the G3 and G5 tetrad planes.  $^1\text{H}$ -NMR spectra at higher temperatures indicated that the quadruplex  $I_T$  is preserved at least up to 65°C, whereas the tetramolecular counterpart  $III_T$  is almost completely destructured at this temperature, as indicated by the disappearance of the imino proton signals (figure 3.16C). Assignment of all aromatic and imino proton signals of  $I_T$  was carried out by mean of a NOESY spectrum (25°C; 300 ms mixing time) and comparison of NOEs with those of  $III_T$  reported by Patel and Hosur<sup>[97]</sup> (data not shown). The pattern of H8/H6–H10–H8/H6 NOE connectivities for  $I_T$  resulted almost completely superimposable to that reported for the corresponding tetramolecular quadruplex  $III_T$ , with the exceptions of G6H8 signal, upfield shifted from 7.75 to 7.66 ppm, and of C7H6 signal downfield shifted from 7.53 to 7.63 ppm. Further, the presence of strong interstrand T4NH–T4CH3 and intrastrand T4CH3–G5NH NOEs confirms that  $I_T$  adopts a T tetrad containing parallel stranded quadruplex structure. The absence of significative chemical shift drifts for aromatic and imino proton signals of  $I_T$  and  $III_T$  demonstrates that the “TEL” linker is capable to stabilize the parallel quadruplex structure of  $I_T$  without any substantial influence on quadruplex topology.  $^1\text{H}$ -NMR spectrum of compound  $II_T$  recorded at 25°C (figure 3.16B) supports the CD data about the aptitude of  $II_T$  to form a parallel quadruplex structure. In fact, the low field region of  $^1\text{H}$ -NMR spectrum of  $II_T$  shows the presence of three strong and two weak partially overlapped signals attributable to exchange protected imino protons of G-quartets. Recording  $^1\text{H}$ -NMR spectra of  $II_T$  at higher temperatures, with 10°C increasing steps, the sharpening and the reduction of intensity of all imino proton signals were observed, followed by their complete disappearance at 75°C. Differently from  $I_T$ , no T-tetrad imino signals were observed in the  $^1\text{H}$ -NMR spectra of  $II$ . The low field region of  $^1\text{H}$ -NMR spectrum of  $IV_T$  recorded at 25°C (figure 3.16D) is characterized by the presence of several H-bonded imino signals. The number of imino signals revealed that more than one quadruplex structure must be present in solution. At higher temperatures, with 10°C increasing steps, the disappearance of some signals and the sharpening of the residual ones were observed. Particularly, five major imino signals were detected at 65°C, four of which closely matched those observed for the parallel-stranded

quadruplex **II<sub>T</sub>** at the same temperature. This observation supports the hypothesis that **4** folds into more than one quadruplex structures. Considering that CD spectra of **IV<sub>T</sub>** showed a positive shoulder band at 289 nm, it is reasonable to hypothesize that **4<sub>T</sub>** in  $K^+$  buffer could fold into a mixture of quadruplex structures, one of which could be of the antiparallel type. If this were the case, it would be the first occurrence of a tetramolecular antiparallel quadruplex structure.



**Figure 3.16:** Aromatic and imino proton region of 500-MHz  $^1\text{H}$ -NMR spectra of **I<sub>T</sub>–IV<sub>T</sub>** (A–D, respectively) in 100 mM  $K^+$  buffer.

The syntheses and investigations of two TEL-ONs, **1<sub>T</sub>** and **2<sub>T</sub>**, characterized by ON sequences potentially able to form T-tetrad containing quadruplexes, have been presented. TEL-ON **1<sub>T</sub>** is based on the sequence TG2TG2C, whose natural counterpart represents, to my knowledge, the first example of an experimentally detected T-tetrad. Collected data from NMR and CD melting experiments clearly indicate that the presence of the tetra-branched linker in **1<sub>T</sub>** not only does not hamper the formation of a T-tetrad, but also provides a higher thermal stability to the complex. The capacity of the tetra-branched linker to stabilize a T-tetrad containing quadruplex structure might be particularly interesting in the research area of quadruplex based aptamers. In fact, the presence of a T-tetrad, thanks to the methyl groups located on the groove surface, could make achievable new hydrophobic interactions capable of improving both the affinity and specificity of a given aptamer toward their target molecule.

## EXPERIMENTAL SESSION

### General Procedures

#### Synthesis and Purification of TEL-ON 1<sub>T</sub> and 2<sub>T</sub>

See above, “ Synthesis and Purification of TEL-ON 1 and 2 ”

The purity of the products was confirmed by HPLC analysis on a HYPERSIL 100-C18 column eluted with a linear gradient CH<sub>3</sub>CN in triethylammonium bicarbonate (TEAB) buffer (pH 7.0, from 0 to 100% in 60 min, flow 1.0 mL/min), with a result higher than 98% (retention time 13.5 and 14.2 min for 1<sub>T</sub> and 2<sub>T</sub>, respectively). The desalifications furnished pure 1<sub>T</sub> and 2<sub>T</sub> (68 and 73 OD<sub>260</sub> units, respectively) which, after lyophilization, were characterized by spectroscopic data.

TEL-ON 1<sub>T</sub>: MALDI-TOF-MS (negative mode) calculated mass: 9070.8; found: 9069.5 TEL-ON 2<sub>T</sub>: MALDI-TOF-MS (negative mode) calculated mass: 9070.8; found: 9069.6

**Preparation of Quadruple Helices(Annealing).** Quadruplexes I<sub>T</sub> –IV<sub>T</sub> were formed by dissolving the TEL-ONs 1<sub>T</sub> –4<sub>T</sub>, in 10 mM K<sub>2</sub>HPO<sub>4</sub> 100 mM KCl, 0.2 mM EDTA (K<sup>+</sup> buffer) and annealed by heating to 90°C for 20 min followed by slow cooling to room temperature. The solutions were equilibrated at 5°C for 24 h before performing the experiments.

**CD Experiments** CD spectra of the quadruplexes were registered on a Jasco 715 circular dichroism spectrophotometer in a 0.1-cm pathlength cuvette at 20°C. The wavelength was varied from 220 to 320 nm at 5 nm min<sup>-1</sup>. The spectra were recorded with a response of 16 s, at 2.0 nm bandwidth, and normalized by subtraction of the background scan with buffer. CD thermal denaturation experiments were performed monitoring the CD value (mdeg) at 265 nm in the range 5–80°C with 1°C/min heating rate. The CD melting experiments were registered at 2.0 · 10<sup>-5</sup> M concentration.

**<sup>1</sup>H-NMR Experiments** NMR data were obtained either on a Varian Unity-INOVA 700 MHz or a Varian Unity-INOVA 500 MHz spectrometer equipped with a broad-band inverse probe with a z-field gradient, and processed using Varian VNMR software package. NMR samples of I–IV were prepared in H<sub>2</sub>O/D<sub>2</sub>O 9:1 at a concentration of single strand of approximately 1.3 mM with a final salt concentration of 100 mM KCl and 10 mM K<sub>2</sub>HPO<sub>4</sub>. One-dimensional (1D) NMR spectra (500MHz) were acquired as 16,384 data points with a recycle delay of 1.0 s at temperatures in the range 25–75°C. Data sets were zero filled to 32,768 points and apodized with a squared shifted sine-bell window function. A 300-ms two-dimensional (2D) NOESY spectrum (700 MHz) was acquired at 25°C as a matrix of 512 - 2048 complex points. Relaxation delay was kept to 1.0 s and a squared shifted sine-bell window function was applied in both dimensions prior to Fourier transformation. A pulsed-field gradient WATERGATE sequence was used for H<sub>2</sub>O suppression in both 1D and 2D spectra.

### 3.11. Final Conclusions

The results, described in this work, confirm the potential of the TEL-ONs, as a new class of ON analogues prone to form intramolecular quadruplex structures, to be more stable than the corresponding tetramolecular counterparts also when a less stable T-tetrad is present in the complex. So the “TEL” approach to the DNA quadruplex can be considered a useful tool for their in vivo applications, furnishing complexes characterized by high stability and predictable structure even at low ON concentrations. The high stability of TEL-structures could be extended to the complexes containing DNA-PNA hybrid strands, to investigate the formation



of PNA guanine quartets, not observed so far. Furthermore, parallel monomolecular TEL-quadruplexes could be used as models in structural studies, being capable to select the formation of only one type of complex, that in some cases is not achievable for tetramolecular complexes. Finally, TEL-ONs show promising properties in terms of rapid kinetic of riassociation so they could be very useful to realize and study nanomotors.

## References

1. W. Guschlbauer, J. F. Chantot, D. Thiele, *J. Biomol. Struct. Dyn.*, **1990**, 8, 491 – 511.
2. M. Gellert, M. N. Lipsett, D. R. Davies, *Proc. Natl. Acad. Sci. USA*, **1962**, 48, 2013 – 2018.
3. E. Henderson, C. C. Hardin, S. K. Walk, I. Tinoco, E. H. Blackburn, *Cell*, **1987**, 51, 899 – 908.
4. J. R. Williamson, M. K. Raghuraman, T. R. Cech, *Cell*, **1989**, 59, 871 – 880.
5. J. SVhnel, *Biopolymers*, **2001**, 56, 32– 51
6. M. Keniry, *Biopolymers*, **2001**, 56, 123 – 146
7. C. G. Hardin, A. G. Perry, K. White, *Biopolymers*, **2001**, 56, 147 – 194.
8. I. Smirnov, R. H. Shafer, *Biopolymers*, **2001**, 56, 209 – 227
9. T. Simonsson, *Biol. Chem.*, **2001**, 382, 621 – 628
10. H. Arthanari, P. H. Bolton, *Chem. Biol.*, **2001**, 8, 221 – 230.
11. J. S. Lindsey, *New J. Chem.*, **1991**, 15, 153 – 180.
12. D. S. Lawrence, T. Jiang, M. Levett, *Chem. Rev.*, **1995**, 95, 2229–2260.
13. S. C. Zimmerman, P. S. Corbin, *Struct. Bonding* (Berlin), **2000**, 96, 63 – 94.
14. T. Kato, *Struct. Bonding* (Berlin), **2000**, 96, 95 – 146.
15. Read, M.; Harrison, R.J.; *PNAS*, **2001**, 98, 4844–4849
16. Luu, K.M.; Phan, A.T.; *J. Am. Chem. Soc.*, **2006**, 128, 9963–9970
17. L. Brunsveld, B. J. B. Folmer, E. W. Meijer, R. P. Sijbesma, *Chem. Rev.*, **2001**, 101, 4071 – 4097.
18. L. J. Prins, D. N. Reinhoudt, P. Timmerman, *Angew. Chem.*, **2001**, 113, 2446 – 2492
19. D. N. Reinhoudt, M. Crego-Calama, *Science*, **2002**, 295, 2403 –2407.
20. Haider, S.M. et al.; *J. Mol. Biol.*, **2003**, 326, 117–125
21. Hardin, C. C.; Henderson, E.; Watson, T.; Prosser, J.K. *Biochemistry* **1991**, 30, 4460–4472.
22. Cech, T. R. *Nature* **1988**, 332, 777–778.
23. Sen, D.; Gilbert, W. *Curr. Opin. Struct. Biol.* **1991**, 1, 435–438.
24. Marathias, V. M.; Bolton, P. H. *Biochemistry* **1999**, 38, 4355–4364.
25. Sen, D.; Gilbert, W. *Methods in Enzymol.* **1992**, 211, 191–199.
26. Wang, Y.; Patel, D. J. *Structure* **1994**, 2, 1141–1156.
27. Marathias, V. M.; Bolton, P. H. *Biochemistry*, **1999**, 38, 4355–4364
28. Sundquist, W. I.; Klug, A. *Nature*, **1989**, 342, 825–829.
29. Kang, C.; Zhang, X.; Ratliff, R.; Moyzis, R.; Rich, A. *Nature* **1992**, 356, 126–131.
30. Wang, Y.; Patel, D. J. *J. Mol. Biol.* **1995**, 251, 76–94
31. Keniry, M. A.; Strahan, G. D.; Owen, E. A.; Shafer, R. H. *Eur. J. Biochem.*, **1995**, 233, 631–643.
32. Kallick, D. A. *J. Am. Chem. Soc.* **1995**, 117, 9941–9950
33. Franklin, S. J.; Barton, J. K. *Biochemistry*, **1998**, 37, 16093–16105
34. Fahlman, R. P.; Sen, D. *J. Mol. Biol.* **1998**, 280, 237–244.
35. Töhl, J.; Eimer, W. *Biophys. Chem.* **1997**, 67, 177–186.
36. Kang, C.; Zhang, X.; Ratliff, R.; Moyzis, R.; Rich, A.; *Nature* **1992**, 356, 126–131.
37. Hud, N. V.; Smith, F. W.; Anet, F. A. L.; Feigon, J.; *Biochemistry* **1996**, 35, 15383–15390.
38. Sen, D.; Gilbert, W. *Biochemistry* **1992**, 31, 65–70
39. Miura, T.; Benevides, J. M.; Thomas, G. J.; *J. Mol. Biol.*, **1995**, 248, 233–238
40. Hardin, C. C.; Henderson, E.; Watson, T.; Prosser, J. K.; *Biochemistry*, **1991**, 30, 4460–4472
41. Boynton, A. L.; McKeehan, W. L.; Whitfield, J.F.; Eds. *Ions, Cell Proliferation and Cancer*; Academic, **1982**.
42. Lau, Y.-T.; Yassin, R.R.; Horowitz, S.B.; *Science*, **1988**, 240, 1321–1323.
43. Roomans, G.; Von Euler, A.; *Cell Biol. Int.*, **1996**, 20, 103–109
44. Rich, A.; *Biochem. Biophys. Acta*, **1958**, 29, 502.
45. O'Brien, E. J.; *Acta. Crystallog.*, **1967**, 23, 92–106.
46. Jing, N.; Hogan, M. E.; *J. Biol. Chem.*, **1998**, 273, 34992–34999.

47. Patel ,P.K.; Hosur, R.V.; *Nucleic Acids Research*, **1999**, 12, 2457-2464
48. Searle, M.S.; Williams, H.E.L.; *Org.Biomol.Chem.*, **2004**, 2, 810-812
49. Famulok ,M.; *Curr. Opin. Chem. Biol.*, **1999**, 9, 324 – 329.
50. Padmanabhan, K.; Tulinsky, A.; *Acta Crystallogr. Sect. D*, **1996**, 52, 272 – 282.
51. Kelly, J. A.; Feigon, J. ; Yeates, T. O.; *J. Mol. Biol.*, **1996**, 256, 417 –422.
52. Vairamani, M.; Gross, M. L.; *J. Am. Chem. Soc.*, **2003**, 125, 42 –43.
53. Jing, N. J.; Rando, R. F. et al.; *Biochemistry*, 1997, 36, 12498 – 12505.
54. Jing, N. J.; Marchand, C.; Liu, J. et al.; *J. Biol. Chem.* 2000, 275, 21 460 – 21467..
55. Wyatt, J. R. Vickers, T. A. et al. ; *Proc. Natl. Acad. Sci. USA* **1994**, 91, 1356 – 1360
56. a) Wen, J.-D.; Gray, D. M.; *Biochemistry* **2002**, 41, 11438 – 11448;  
b) Wen J.-D., Gray, C.W. ; Gray, D. M; *Biochemistry*, **2001**, 40, 9300 – 9310;  
c) Tarrago-Litvak, L.; Litvak, S.; Andreola, M. L.; *J. Mol. Biol.*, **2002**, 324, 195 – 203.
57. Dapic, V.; Bates, P. J. et al.; *Biochemistry* , **2002**, 41, 3676 – 3685
58. Shaffitzel, C.; Berger, I.; Postberg, J., Hanes, J.;*Proc. Natl. Acad. Sci. USA*, **2001**, 98, 8572– 8577.
59. Padmanabhan, K.; Tulinsky, A.*Acta Crystallogr. Sect. D*, **1996**, 52, 272 – 282.
60. Huizenga, D. E.; Szostak, J.W.; *Biochemistry*, **1995**, 34, 656 – 665
61. a) Han, H. Y.; Langley, D. R.; Rangan, A.; Hurley, L. H.; *J. Am.Chem. Soc.*, **2001**, 123, 6485 – 6495 ;  
b) Shi, D. F.; Wheelhouse, R. T.; Sun, D. Y.; Hurley, L. H. ;*J. Med. Chem.*, **2001**, 44, 4509 – 4523.
62. Travascio, P.; Witting, P. K.; Mauk, A. G.; Sen, D.; *J. Am. Chem.Soc.*, **2001**, 123, 1337 – 1348.
63. Isalan, M.; Patel, S. D; Balasubramanian, S.; Choo, Y.; *Biochemistry*, **2001**, 40, 830 – 836.
64. Guo, Q.; Lu, M.;Marky; L. A.; Kallenbach, N. R, *Biochemistry* **1992**, 31, 2451 – 2455.
65. Sun, D. Y.; Thompson, B., Cathers, B. E. et al.; *Med. Chem.*, **1997**, 40, 2113 – 2116.
66. Gavathiotis, E.; Heald; R. A. Stevens, M. F. G.; Searle, M.S.; *Angew. Chem.*, **2001**, 113, 4885 – 4887;  
*Angew. Chem., Int. Ed.*, **2001**, 40, 4749 – 4751.
67. Read, M. A.; Neidle, S.; *Biochemistry*, **2000**, 39, 13422 – 13432.
68. Haider, S. M.; Parkinson, G. N; Neidle, S.; *J. Mol. Biol.*, **2003**, 326, 117 – 125.
69. Schouten, J. A.; Ladame, S. Mason, S. J. et al.; *J. Am. Chem. Soc.*, **2003**, 125, ASAP April 17.
70. Fedoroff Yu. O.; Salazar, M.; Han, H. et al.; *Biochemistry*, **1998**, 37, 12367 – 12374.
71. Koeppe, F.; Riou, J. F.; Laoui A, et al.; *Nucleic Acids Res.* **2001**, 29, 1087 – 1096.
72. Shin-ya, K.; Wierzba, K.; Matsuo, K. et al., *J. Am. Chem. Soc.* **2001**, 123, 1262 – 1263.
73. Kim, M. Y. ; Vankayalapati, H.; Kazuo, S. et al.; *J. Am. Chem. Soc.*, **2002**, 124, 2098 – 2099.
74. Zimmerman, S. B.; *J. Mol. Biol.*, **1976**, 106, 663 – 672.
75. Marotta, S. P.; Tamburri, P. A.; Sheardy R. D.; *Biochemistry*, **1996**, 35, 10484 – 10492.
76. Marsh, T. C.; Henderson, E.; *Biochemistry*, **1994**, 33, 10718 –10724.
77. Charles, J. A. M.; McGown, L. B., *Electrophoresis*, **2002**, 23, 1599 – 1604.
78. Nutiu, R.; Li, Y. F; *J. Am. Chem. Soc.*, **2003**, 125, 4771 – 4778.
79. Mergny, J.-L.; Maurizot, J.-C.; *ChemBioChem* **2001**, 2 124 – 132.
80. Alberti, P.; Mergny, J.-L; *Proc. Natl. Acad. Sci. USA* **2003**, 100, 1569 – 1573.
81. Li, J. W. J.; Tan, W. H.; *Nano Lett.*,**2002**, 2, 315 – 318.
82. Aviram, A. ; Ratner M.; *Chem. Phys. Lett.*, **1974**, 29, 277 – 283
83. Joachim, C.; Gimzewski, J. K ; Aviram A.; *Nature*, **2000**, 408, 541 – 548.
84. Porath, D. Bezryadin, A. de Vries S., Dekker, C. ; *Nature*, **2000**, 403, 635 – 638.
85. a) Haider, S. M.; Parkinson, G. N.; Neidle, S.;*J. Mol. Biol.* **2003**, 326,117 – 125;  
b) Clark, G. R.; Pytel, P. D.; Squire, C. J.; Neidle, S.; *J. Am. Chem.Soc.*, **2003**, 125, 4066 – 4067.
86. Mergny, J.L., De Cian, A.; Ghelab, A.; *Nucleic Acids Research*, **2005**, 33, 81-94
87. Borbone, N.; Oliviero, G. et al; *Tetrahedron Lett.*, **2004**, 45, 4869-4872
88. Amato, J.; Borbone, N., et al.; *Bioconjugate Chemistry*, **2006**, 17, 889-898
89. Amato, J.; Oliviero, G., et al; *Biopolymers*, **2006**, 81, 194-201
90. Patel, P.K.; Hosur, R.V.; *Nucleic Acids Research*, **1999**, 12, 2457-2464

## Chapter 4

### ***“Synthesis of combinatorial libraries of nucleosides, specifically inosine analogues”***

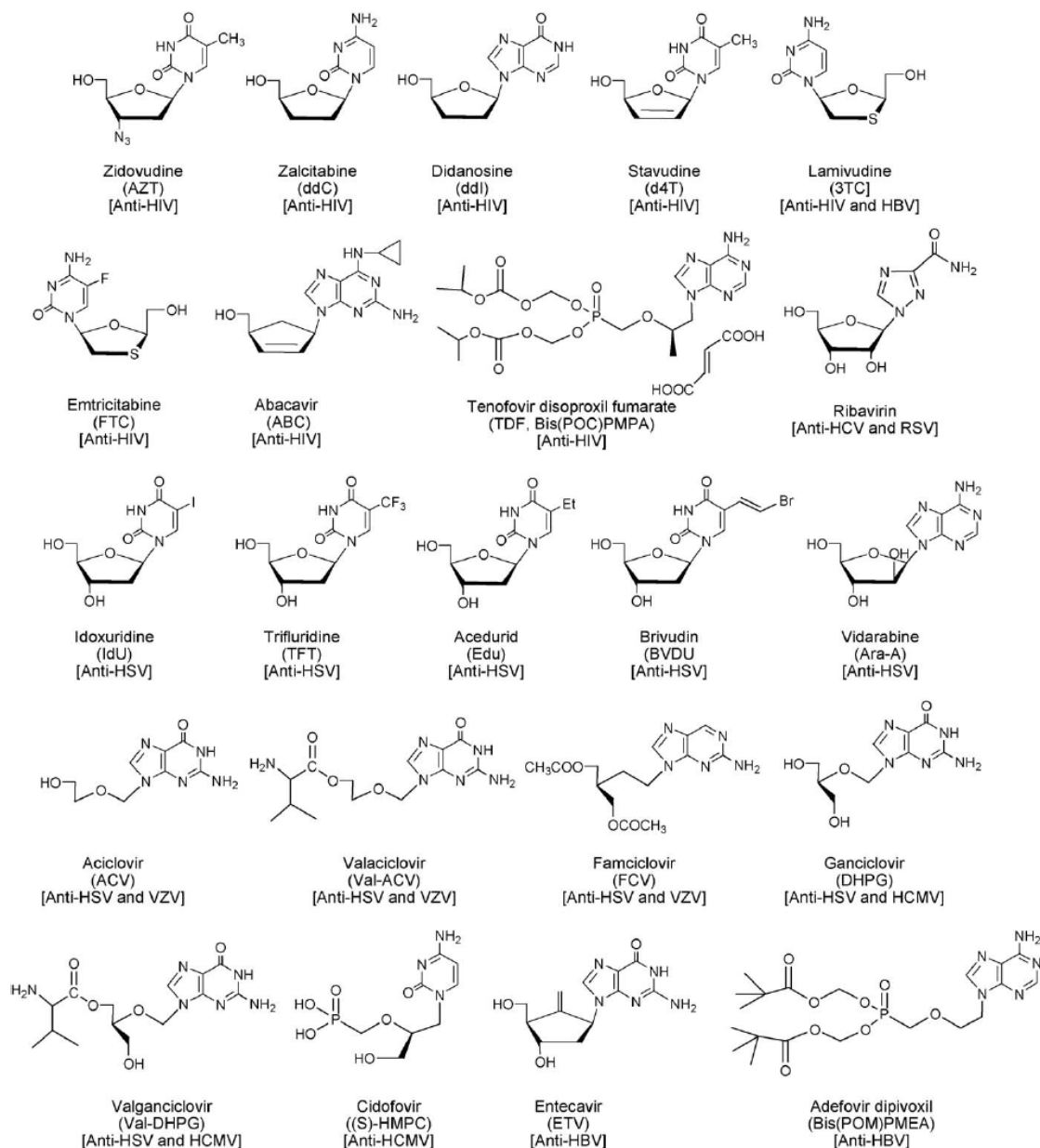
#### ***Introduction***

Nucleoside research has historically been an area of keen interest for small molecule drug discovery due predominately to the medicinal value and biological importance of these molecules. Apart from being the genomic building blocks, nucleosides interact with roughly one-third of the protein classes in the human genome, including polymerases, kinases, reductases, motor proteins, membrane receptors, and structural proteins, all targets of biological importance.

The binding motifs of these nucleosides are associated with a broad array of targets of therapeutic importance in biological systems. It was early recognized that, introducing diversity into the carbohydrate or the base subunits of nucleosides, represents promising strategies to identify specific receptor ligands, enzyme inhibitors, or nucleoside function modifiers. Nucleoside analogues, showing very complex and interesting structures, have been isolated from t-RNA (5-methylcytidine, 3'-methyluridine), marine organism and several other microbiological forms. Naturally occurring nucleosides analogues demonstrate selective activities, such as protein synthesis inhibitors (puromycin), glycosyl transferase inhibitors (tunicamycin), methyl transferase inhibitors (sinefungin) and antibiotic activity (coformicine A and adenomicine). The discovery that some nucleoside analogues can possess biological activities has been a significant breakthrough especially in antiviral and anticancer chemotherapy. In this regard, the isolation from natural sources, but especially synthetic work have lead to the discovery of a large variety of new analogues. After decades of drug discovery and development effort from hundreds of academic laboratories and pharmaceutical companies, approximately 50-nucleoside- and nucleotide-related drugs, have been approved by the US Food and Drug Administration, and pushed to the market, and over 80 nucleosides/nucleotides are in the (pre) clinical studies for various therapeutic indications, including antiviral<sup>[1]</sup>, anticancer<sup>[2]</sup>, antibiotic<sup>[3]</sup> and antifungal<sup>[4]</sup>.

In the search for antiviral nucleoside analogues, structural modifications of the heterocyclic bases and/or modifications on the sugar moiety of natural nucleosides, can be attempted. In the latter, the main modifications involved changes in the (2-deoxy)-D-ribofuranose moiety like, inversion of hydroxyl group configurations, their substitution/functionalisation by various synthetic groups, or cleavage of the sugar ring leading to acyclic nucleosides. Other structural modifications have also been attempted such as replacement by a methylene group or a sulfur atom of the endocyclic oxygen, transposition of the latter and/or additional insertion of a second heteroatom in the sugar moiety. Currently, nucleoside analogues are prominent drugs in the management of several viral infections. The nucleoside analogues at present formally approved for the treatment of viral infections are shown in figure 4.1 Acyclovir is an important antiviral in hepatitis C (HCV) treatment, while Lamivudine (3'-thiocytydine, 3TC) is a nucleoside analogue used for treatment of hepatitis B virus (HBV), the 5-iodouridine is employed in Herpes Simplex infections (HSV). Furthermore, it must be considered that the most commonly used anti-HIV drugs are nucleoside analogues, including Zidovudine (AZT, azidothymidine), Hivid (ddC, zalcitabine), Zerit (d4T, stavudine) and some others, all possessing high activity in antiretroviral therapies<sup>[4]</sup>. They mimic natural nucleosides and

target a unique, but complex viral polymerase that is essential for viral replication. Some nucleosides show also antiproliferative properties, for example arabinosylcytidine and 5-azacytidine are both utilized in the treatment of leukaemia, while 5-fluorouracil (5-FU) is employed in breast tumour.



**Figure 4.1:** Nucleoside analogues currently used in antiviral therapy.

For salvage to be efficient, mediated transport of nucleosides and nucleobases across the plasma membrane has to occur. Then metabolic activation of nucleosides and nucleobases takes place, modulating their pharmacological action. The biochemical mechanism can be of various type, but in most cases the 5'-phosphorylation of the sugar moiety is the first step. The nucleoside 5'-phosphate is then converted into a triphosphate derivative that inhibits, in different way, the replication of DNA or RNA. However, some drugs such as AZT, ddI, 3TC etc. rapidly develop drug resistance and show mitochondrial, bone-marrow and other toxicity. The increasing resistance of pathogens, the often severe side effects of nucleosides in chemotherapy, and the lack of selective ligands for

adenosine receptor subclasses, despite extensive medicinal chemistry research, emphasizes the need for nucleoside analogues in high number and diversity. In addition, the availability of high-throughput screening capabilities, together with the combinatorial synthesis of small organic molecule libraries, offers a unique opportunity to accelerate the discovery of novel pharmaceutical targets and leads, especially, with biologically privileged scaffolds, such as nucleosides in hand.

#### 4.1. Nucleoside analogues

The design of selective, specific and non-toxic antiviral, but also anticancer agents, has presented extraordinary challenges compared with the design of other antimicrobial agents. This is primarily because, in contrast to bacteria and parasites that are equipped with their own metabolic machinery for reproduction and growth, viruses rely on the host cellular machinery for replication and propagation. Consequently, there are few virus-specific molecular targets that are amenable to antiviral intervention. In addition, drug resistance to antiviral agents is a major problem as many viruses mutate rapidly under the selective pressure of the antiviral therapy. In recent times, however, and largely owing to the AIDS epidemic and advances in molecular virology, the structure and function of a few virus-specific molecular targets such as proteases and helicases have been revealed. Furthermore, in the past decade, a repertoire of discovery tools in terms of structure and mechanism-based drug design, such as elucidations on *nucleoside analogues uptake and export, their activation, mechanism of action and mitochondrial toxicity*, have provided novel strategies for the design of antiviral agents against new molecular targets. In this regard different sites of diversity on both purine- and pyrimidine-nucleoside scaffolds have been introduced, schematically represented in figure 4.2, bearing a myriad of functional groups. Generally, nucleosides analogues can be catalogued in “**Base-modified**” and “**Sugar-modified**” nucleoside analogues.

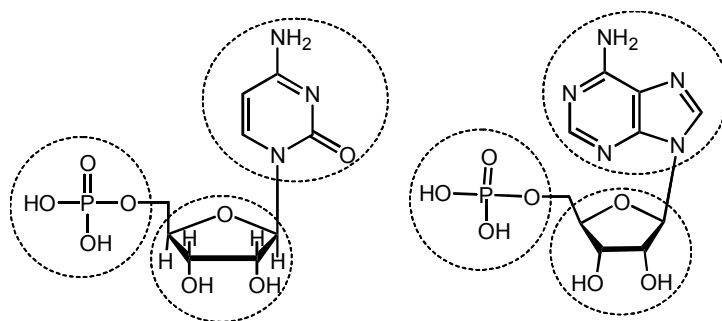
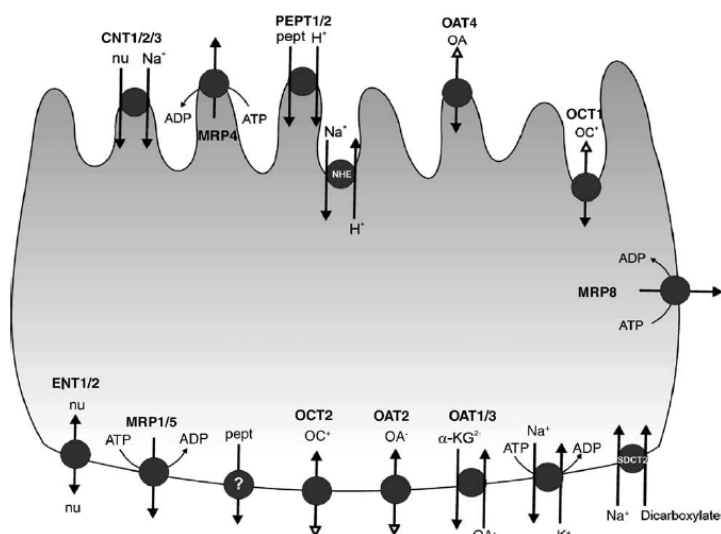


Figure 4.2: Nucleoside, nucleotide modifications

Notwithstanding these advances, progress in the discovery of new antiviral therapeutics has remained slow, and current antiviral chemotherapy is heavily dependent on a “cocktail” of antiviral agents. New antiviral agents, with unique mechanisms of action, are urgently needed, given the propensity of viruses to mutate rapidly under the selective pressure of chemotherapy. In the last years, the combinatorial chemistry has emerged as very efficient tool to produce a large number of new molecules differing in punctiform manner around a core structure of active compounds. Solid phase combinatorial strategies have been mainly used in the nucleoside field.

## 4.2. Nucleoside analogue uptake

The intracellular concentration of nucleosides and nucleobases is of great importance for the success of antiviral and anticancer therapies and, to some extent, it is the result of the metabolic background of the specific cell line used for infection studies, its particular suit of enzymes and transporters. Transporter-mediated pathways are involved in either the uptake or the efflux of nucleoside and nucleobase derivatives. From a biochemical point of view, four different types of transport processes for nucleoside-related drugs are involved: 1) equilibrative uniport; 2) substrate exchange; 3) concentrative  $\text{Na}^+$  - or  $\text{H}^+$ -dependent uptake and finally 4) substrate export through primary ATP-dependent active efflux pump. These mechanisms are mainly related to the following set of transporter families: Concentrative Nucleoside Transporter (CNT), Equilibrative Nucleoside Transporter (ENT), Organic Anion Transporter (OAT) and Organic Cation Transporter (OCT), Peptide Transporter (PEPT) and Multidrug Resistance Protein (MRP)<sup>[5]</sup>. The scheme in figure 4.3 shows an idealistic epithelial cell expressing all known nucleoside-based antiviral drug transporters, emphasizing their location, either apical or basolateral, which ensure, when necessary, vectorial flux of substrates across the epithelium. Although information regarding the plasma membrane domains, in which they are inserted, has recently become available, the location for some specific transporters, such as Multidrug Resistance Proteins (MRP8) is still in doubt and so they are shown at the lateral side of the cell.



**Figure 4.3:** Major transporters implicated in the uptake and efflux of nucleoside-derived drugs used in antiviral therapy. A model of an idealized polarized epithelial cell, with the apical and basolateral membrane domains.

Equilibrative uniport works down the substrate concentration gradient and thus, it is in principle a reversible mechanism that responds only to the transmembrane substrate gradient tending to equilibrate it. This would be the case for the basolateral Equilibrative Nucleoside Transporter (ENT) carriers. The working model for an exchanger, particularly an obligatory exchanger, implies that efficient translocation of the substrate occurs only if both molecules are bound at both sides of the membrane to allow substrate exchange. An example of this mechanism depicted in figure 4.3 is that the Organic Anion Transporter, also located at the basolateral plasma membrane, which exchanges an extracellular organic anion with a dicarboxylic acid, such as alphaketoglutarate. This process may be thermodynamically favourable, due to the ability of certain cells to concentrate

dicarboxylates thanks to a complementary transporter, such as SDCT2 (also shown in Figure 4.3), which allows the cell to concentrate  $\alpha$ -ketoglutarate in a  $\text{Na}^+$ -dependent manner. The latter is an example of the fourth mechanism of transport processes listed above. Coupling with  $\text{Na}^+$  also occurs for other transporters, such as Concentrative Nucleoside Transporters (CNTs), implicated in the uptake of natural nucleosides, whereas  $\text{H}^+$ -dependent uptake is involved in Peptide Transporters (PEPTs). These two basic mechanisms of concentrative transport,  $\text{Na}^+$ - and  $\text{H}^+$ -dependent, require tight coupling with carriers responsible for the immediate recovery of the ionic transmembrane gradients, the  $\text{Na}^+$ ,  $\text{K}^+$ -ATPase and the  $\text{Na}^+/\text{H}^+$  exchanger. In contrast to these secondary and tertiary active transport mechanisms, ATP-dependent primary export pumps, responsible for antiviral drug efflux, have recently been identified. This would be the case of the Multidrug Resistance Protein isoforms (MRPs), shown in Figure 4.3, which mediate efflux, also in an energy-dependent manner, against substrate concentration, thanks to a direct coupling with ATP hydrolysis. The hypothesis that transporters belonging to the same gene family are responsible for the uptake of structurally related compounds is valid, at least for those responsible for uptake processes other than export pumps such as MRPs. Nevertheless, structurally related molecules can often be substrates of completely unrelated carrier proteins. However, when natural substrates are modified for pharmacological purposes, the spectrum of transporter proteins able to recognize them may vary.

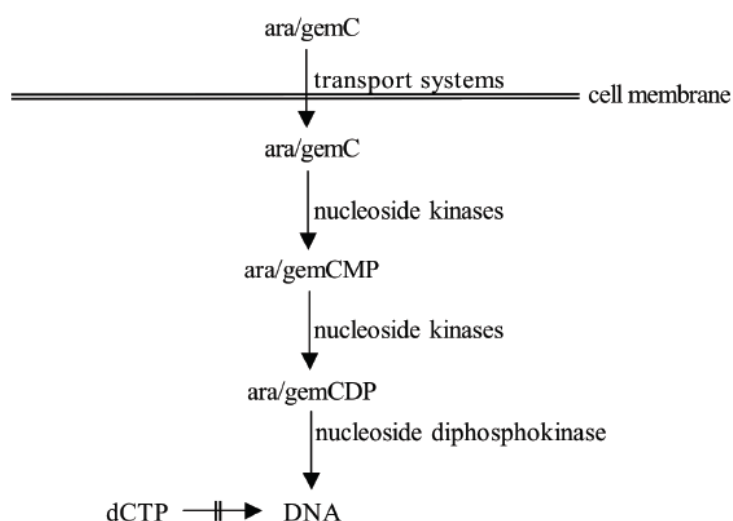
#### **4.3. Nucleoside analogue activation**

Phosphorylation is required for mutagenesis and nucleotide metabolism interference. Nucleosides are phosphorylated by intracellular kinases to their active fraudulent nucleoside triphosphate (TP) analogue form, which compete with the natural dNTP for incorporation into the elongating proviral DNA. Since all current nucleoside-TP analogues lack a 3'-hydroxyl group, incorporation results in chain termination. The fact that a specific nucleoside-TP analogue competes with its natural nucleotide each time, and also it appears in the active site near the reverse-transcribing viral RNA chain, produces multiple opportunities for their blockade during a single round of reverse transcription. The long intracellular half-life of many nucleoside-TP analogues can result in a significant post-antibiotic effect<sup>[6]</sup>. Factors that influence their efficacy and dose frequency include the drug concentration versus time profile in plasma (pharmacokinetics), the efficiency of cellular uptake of the nucleoside analogues, levels of phosphorylation to their correspondent TP-derivatives, and their inhibition constant ( $K_i$ ) with the target (viral DNA-polymerases, reverse transcriptase, respectively in DNA and RNA viruses).

#### **4.4. Mechanism of action of nucleoside analogues**

Although all orally available drugs target the viral polymerase, nucleoside analogues have different mechanisms of action on the viral genome replication machinery. Several drugs such as adefovir, entecavir and amdoxovir inhibit the priming of reverse transcription, a unique enzymic reaction that results in the synthesis of a short DNA primer covalently attached to a conserved tyrosine residue of the viral polymerase. Other compounds inhibit the elongation of viral minus-strand DNA, such as lamivudine, emtricitabine and elvucitabine. The mechanism of action of clevudine involves a weak effect on the priming reaction but a stronger inhibitory activity on plus-strand DNA synthesis. The mechanism of action of telbivudine has not been described so far. Based on these distinct modes of action, one may speculate that the combination of drugs targeting different steps of viral

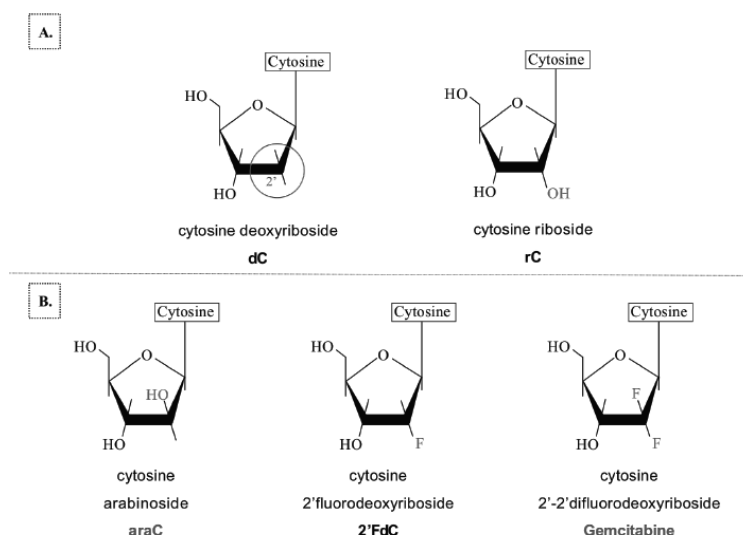
genome replication may lead to an additive or synergic effect<sup>[7]</sup>. Therefore, the evaluation of the combination of drugs such asemtricitabine and clevudine, which are under clinical investigation, is relevant for the future development of combination therapy. Nucleoside analogues are also an emerging field of novel anticancer drugs, that is revolutionizing cancer therapy. In addition to inhibiting DNA replication by one or more mechanisms, these compounds may produce effects subsequent to their incorporation into the DNA; for example, premature DNA chain termination, structural lesions in the DNA product, and effects on transcription of modified DNA. Since charged nucleotides were believed normally do not enter the cells and considered ineffective as drugs, nucleosides have proved critical as therapeutic agents. Nucleosides and their analogues are precursors, that are intracellularly phosphorylated ('activated') into nucleotides (dNTPs), and act to inhibit or are incorporated by DNA polymerases *in vivo* (figure 4.4).



**Figure 4.4:** The intracellular formation of nucleotide triphosphate for araC and gemcitabine (gemC), as well as, their incorporation into new DNA.

Some two of these nucleoside analogues are currently used in chemotherapy or are in clinical trials or development. Examples include arabinofuranosylcytosine (araC)<sup>[8]</sup> and 5-FU (the first for the treatment of non-solid tumors like leukemia, while the latter involved in breast tumour treatment), and 2',2'-difluorodeoxycytosine (gemcitabine)<sup>[9]</sup> (for treatment of various non-solid and solid tumours, figure 4.5). Promising applications of nucleotides analogues in antiviral and anticancer therapy are currently evaluated.





**Figure 4.5:** Nucleoside structures highlighting the differences on the groups on the 2'-carbon of the ribose-ring. **A.** Naturally occurring nucleosides dCTP and rCTP, and **B.** sugar-modified nucleoside analogues, araC, 2'FdC and gemcitabine.

#### 4.5. Mitochondrial toxicity of nucleoside analogues<sup>[6]</sup>

Generally, the mechanism of action of nucleoside analogues is based upon the intracellular phosphorylation to their 5'-triphosphate form which can interact with virus-specific polymerases, acting as a competitive inhibitor or an alternate substrate for these target enzymes, usually preventing further viral nucleic acid chain elongation. Long-term toxicities associated with nucleosides phosphorylation may be related to over-activation of this process. Elevated cell activation results in high nucleic acid synthesis and an up-regulation of kinases that phosphorylate nucleosides. Most of the clinical manifestations, resemble mitochondrial disease, and histological evidence, demonstrates abnormal mitochondria (mt) and/or mtDNA depletion in affected tissues. Studies show that nucleoside tri-phosphates competitively inhibit mtDNA polymerase gamma in vitro, the sole enzyme responsible for the base excision repair of oxidative damage to mitochondrial DNA. This in turn may decrease the number of mitochondrial respiratory chain proteins, inhibit aerobic respiration, induce oxidative stress, increase mutation in mtDNA, and result in mitochondrial and/or tissue failure, evidenced by the toxic accumulation of non-esterified fatty acids, dicarboxylic acids and free radical damage.

#### 4.6. Nucleotide analogues

Since the discovery that monophosphorylated forms of nucleosides can be carried in cells by the Multidrug Resistance Proteins (MRP), and known that, monophosphorylated (MP) derivatives show smallest drug-induced cytotoxicity, AZT MP (Schuetz et al. 1999), d4T MP (Reid et al., 2003) and 5-FU MP (Guo et al., 2003) have been synthesized and evaluated in MRP uptake, in their activities and toxicities. Adefovir dipivoxil (ADV), the first prototype nucleotide analogue, has been approved recently for HBV therapy<sup>[10]</sup>. ADV seems relatively non-toxic upon longer-term use, overcome viral resistance to chemotherapy, and prevent viral replication rebound following cessation of therapy. Towards this objective, there is genuine optimism that selected nucleotide analogues could become front-line agents to be used in combination with nucleosides and other agents in antiviral therapy producing improved pharmacological effects. Furthermore,

nucleotide analogues can be useful tools in investigating the DNA polymerase mechanism. Considering the irreplaceable role of DNA polymerases in numerous important molecular biological core technologies, future potentials for biotechnological applications could be envisaged for nucleotide analogues. However, all these promising applications are limited by the broad tissue distribution of MRP proteins, including endothelia, most epithelia, smooth muscle and immune system cells (Adachi et al.,2002). However, the list of tissues and/or cell types is increasing, although there is discrepancy between the detection of the mRNA and the corresponding protein. Unfortunately, although the pharmacology of these transporter proteins is being studied, there is little information about the regulation of MRP activity, with the exception of the evidence that specific isoforms can be over-expressed in particular tumours.

#### **4.7. Solid-phase synthesis of nucleoside analogues**

A variety of different modified nucleosides derivatives, above reported, were synthesized by the classical approaches in solution, which is expensive and time consuming. Combinatorial synthesis of large diverse libraries and high-throughput screening technologies have recently emerged as powerful drug discovery paradigms. Various solid-phase<sup>[11]</sup>, solution-phase<sup>[11]</sup>, liquid phase<sup>[12]</sup>, and third-phase combinatorial approaches have been successfully utilized for the generation of different oligomeric and small molecule libraries for a wide range of biological screenings. Unfortunately, these power technologies, especially parallel solid-phase combinatorial strategies<sup>[13]</sup>, have not been applied to the nucleoside chemistry for small molecule drug discovery yet although di/tri- and oligonucleotides<sup>[14]</sup> were synthesized on solid support, and the modified solid supports have been used as acylating agent to acylate nucleosides derivatives<sup>[15]</sup>. Several key considerations had to be taken into account in order to develop a strategy for the combinatorial solid-phase synthesis of nucleosides analogues in high number and short period of time, to explore wide biological activities: i) most applicable linkers do not meet the general requirements for the solid phase synthesis of nucleoside libraries: stable enough under the required reaction conditions during synthesis, and labile enough to be cleaved easily from solid support without affecting nucleoside products; ii) limited positions on nucleoside can be attached on solid support; iii) limited types of reactions can be utilized and limited number of sites can be utilized and limited number of sites can be combinatorialized on the nucleoside skeletons, which prevent the generation of large diverse nucleoside libraries; iv) it is more difficult to get high quality and purity nucleoside libraries without purification compared to other small molecule libraries. Only in the most recent years, a variety of solid phase combinatorial strategies have been reported for the preparation of nucleoside and small oligonucleotide analogues libraries. In these approaches both the chemical stability and the position of the linkage with the polymeric support, play an important role in the selection of the chemical treatments allowing the nucleobase or sugar-phosphate modifications. For example, a solid support binding uridine or thymidine nucleosides by the acid labile 5'-O-trityl linkage has been used to prepare libraries of N-4-alkylated cytidine derivatives<sup>[16]</sup>. Another recently reported nucleoside functionalized support uses an alkaline labile 5'-O-succinyl linkage to bind the 6-chloro-2-nitro inosine derivative to produce a set of N-6- or N-2 alkylated adenosine derivatives<sup>[17]</sup>. Also the 2',3'-acetal linkage has been used to bind purine and pyrimidine nucleosides to a solid support which has employed to introduce chemical diversities both on the base (C-6 of purine and C-4 of pyrimidine) and on the 5'-ribose position<sup>[18]</sup>. In another approach, a pyrimidine or purine nucleoside is anchored to the polymeric support by the N-3 or N-1 base position, respectively, through a N-alkyl- $\beta$ -thioether function<sup>[19-21]</sup>,

which results stable to both the acidic and basic conditions. The above supports have been used to prepare 2', 3'-ribose modified nucleoside or nucleotide analogues. Furthermore, CPG support loading 5'-DMT-nucleosides by the classical 3'-succinyl linkage<sup>[22-23]</sup> or by a 3' acyloxyaryl phosphate linker<sup>[24]</sup>, have been successful exploited to prepare very large nucleic-acid-bases (NAB<sup>TM</sup>) libraries of nucleosidic 5'-phosphoramidate derivatives, and nucleic acid fragments to be tested in their antiviral activity<sup>[25]</sup>.

#### 4.8. The aim of the work

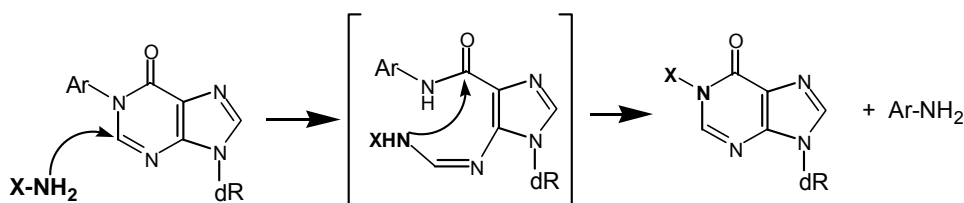
Although hundreds of purine and pyrimidine nucleosides were synthesized for a variety of biological and biomedical studies only a handful of N1-alkylated-inosine derivatives and the corresponding 2',3'- seco-nucleosides were reported.

To thoroughly explore the biological and biomedical properties of these analogues, the design and synthesis of these derivatives were performed by combinatorializing the N1-(2,4-dinitro)-phenyl derivatized position with a variety of amino building blocks, obtaining a first small library of N1-alkylated inosine and AICAR derivatives.

Moreover, the supports bearing N1-alkylated inosine and AICAR derivatives, furnished, by a solid phase cleavage of the 2', 3' ribose bond, a set of new N1-alkylated-2',3'-secoinosine derivatives and AICAR derivative ones, in high yields.

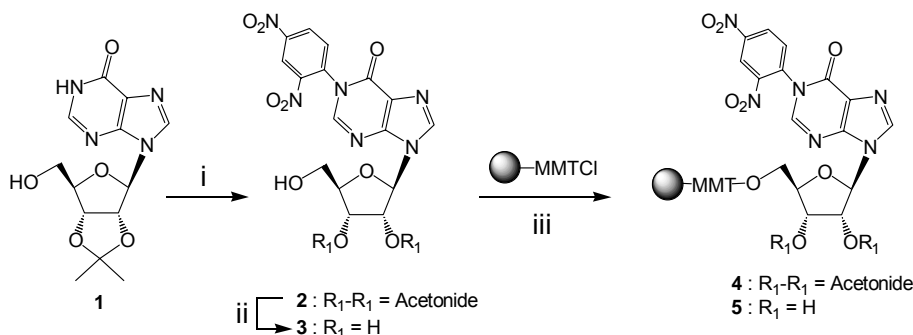
#### 4.9. Chemistry

In an effort to enlarge the nucleoside chemical reactivity on the solid phase, and thus the number of obtainable structurally diverse analogues, the synthesis and exploitation of the new acid labile nucleoside functionalized supports **4** and **5**, which bind the N-1-dinitrophenyl-inosine derivatives **2** or **3** through the ribose 5'-position, is reported. These supports have been employed in the solid phase synthesis of the N-1 substituted inosine **8a-e**, **9a-e**, the related 2', 3'-seconucleoside derivatives **11a-e** and the AICAR derivatives **14** and **15**. The solid phase strategy, here proposed, is based on our previous studies on the C-2 reactivity of N-1-dinitrophenyl-2'-deoxy-inosine towards N-nucleophiles<sup>[26-27]</sup>, to obtain N-1 substituted inosine and AICAR derivatives. The reported reaction mechanism (figure 4.6) indicates that, when a strong electron-withdrawing group, such as the 2,4-dinitrophenyl (DNP), nitro group<sup>[28]</sup> (NO<sub>2</sub>) or arylsulfonyl<sup>[29]</sup> one (ArS) is attached to the N-1 atom of the hypoxanthine ring, the C-2 carbon become electrophilic enough to react with amino nucleophiles (R-NH<sub>2</sub>), leading to N1 substituted inosine derivative, by a fast opening and re-closure of the six terms purine cycle. It is to be noted that, as a consequence of the purine rearrangement, the endocyclic N-1 atom is substituted by the nitrogen atom of the nucleophilic reactant. This purine reactivity has been also exploited by others to introduce a modified base into oligonucleotides<sup>[30]</sup>.



**Figure 4.6:** The nitrogen of nucleophile substitutes the purine N-1 nitrogen.

To exploit this reaction in a solid phase strategy to obtain a small library of N-1 alkyl inosine derivatives, we bound the 1-(2,4-dinitrophenyl)-2',3'-O-isopropylideneinosine **2**, or the corresponding unprotected inosine derivative **3** to the commercially available polystyrenemonomethoxytrityl chloride (MMTCl) resin by 5'-O-trityl ether linkage (iii, Scheme 4.1).

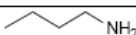
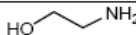

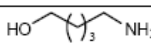
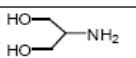
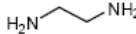


**Scheme 4.1:** Acid labile N-1-(2,4-dinitrophenyl)-nucleoside functionalized supports **4** and **5**

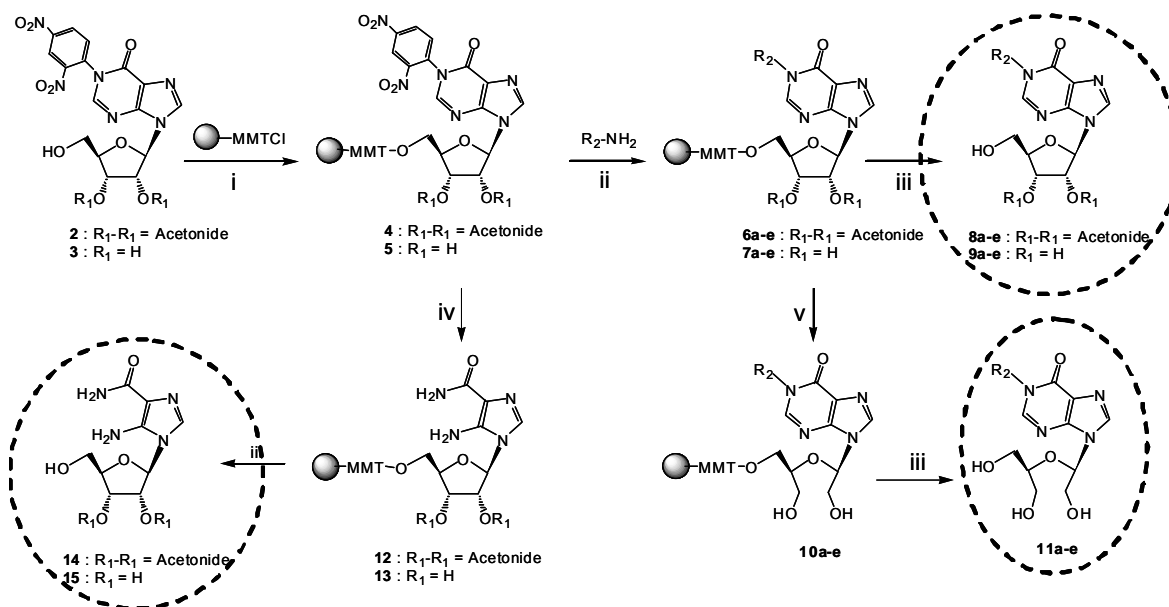
Inosine derivative **2** was synthesized by the reaction of the commercially available 2',3'-O-isopropylidene inosine (**1**) with 2,4-dinitrochlorobenzene (DNCIB), essentially as previously described<sup>[31]</sup>(i, Scheme 4.1). The 2',3'-deprotected-inosine-derivative **3**, was obtained treating **2** with aqueous formic acid (5 h, r.t., 90 % yield, Scheme 4.1, ii). The reaction of the MMTCl polystyrene resin (1.3 meq/g) with **2** or **3** (2 equiv.) in anhydrous pyridine at room temperature and in presence of 4-(N,N-dimethylamino)-pyridine (DMAP), afforded the support **4** or **5** respectively, in almost quantitative yields (iii, Scheme 4.1). The structure and the loading of the supports **4** and **5** were confirmed analyzing, by NMR and quantitative UV experiments, the released inosine derivatives **2** and **3**, respectively, by treatment with 2% TFA in DCM (8 min r.t.).

Support **4** and **5** were then reacted with several N-nucleophiles ( $\text{R}_2\text{-NH}_2$ , Table 4.1 entry **a-e**) to give supports **6a-e** and **7a-e**, respectively (scheme 4.2). In a typical reaction, 100 mg (0.13 mmol) of support **4** (or **5**), swollen in DMF, was left in contact with the  $\text{R}_2\text{-NH}_2$  nucleophile (5.0 mmol) in 1.5 mL of DMF under shaking (8 h, at 50 °C). After washings with DMF and MeOH, the support was dried under reduced pressure and the reaction yield was evaluated detaching the nucleoside material from a weighted amount of resin. The reaction of **4** or **5** with ethylenediamine (Table 4.1, entry **f**) furnished, as expected, the supports **12** and **13** bearing 2'-3'-isopropylidene-AICAR and AICAR, respectively, in almost quantitative yields. The structure of the support **6**, **7**, **12**, and **13** were ascertained analyzing and purifying by HPLC,  $^1\text{H}$  NMR and MS analyses the corresponding detached crude materials **8**, **9**, **14**, and **15**. The product yields and the  $^1\text{H}$  NMR data are reported in Table 4.1.

**Table 4.1:** Reactions of the support **2** Products **4**, **6**, and **8**. N.T: not tested<sup>a</sup> Starting from resin **2**. <sup>b</sup> 400 MHz, (CD<sub>3</sub>OD) significant protons at ppm. <sup>c</sup> Starting from resin **3**.

Entry	R <sub>2</sub> -NH <sub>2</sub>	8, 9, 14, 15 Yield <sup>a</sup> (%)	9a-e and 15 <sup>1</sup> H NMR <sup>b</sup>		11a-e Yield <sup>c</sup> (%)	<sup>1</sup> H NMR <sup>b</sup>	
			H-2; H-8; H-1'	R <sub>2</sub> moiety		H-2; H-8; H-1' (seco)	R <sub>2</sub> moiety
<b>a</b>		<b>8</b> (98)	8.41; 8.36;	4.12; 1.76;	<b>11a</b> (85)	8.30; 8.26;	4.00; 1.75;
		<b>9</b> (98)	6.02	1.40; 0.98		6.04	1.38; 0.96
<b>b</b>		<b>8</b> (96)	8.42; 8.26;	4.19; 3.82	<b>11b</b> (85)	8.28; 8.24;	4.20; 3.83
		<b>9</b> (95)	6.01			6.04	
<b>c</b>		<b>8</b> (98)	8.41; 8.32;	4.20; 3.60;	<b>11c</b> (84)	8.31; 8.26;	4.21; 3.60;
		<b>9</b> (96)	6.02	1.98		6.03	1.98
<b>d</b>		<b>8</b> (98)	8.40; 8.35;	4.12; 3.56	<b>11d</b> (82)	8.33; 8.27;	4.11; 3.55;
		<b>9</b> (98)	6.02	1.81; 1.59; 1.42		6.03	1.79; 1.58; 1.43
<b>e</b>		<b>8</b> (92)	8.38; 8.32;	3.98 (2CH <sub>2</sub> OH);	<b>11e</b> (75)	8.29; 8.04;	4.02 (2CH <sub>2</sub> OH)
		<b>9</b> (90)	6.00	3.92 (CH)		6.05	3.95 (CH)
<b>f</b>		<b>1</b> (98)	8.04;		N.T.	N.T.	
		<b>1</b> (98)	5.66				

The second goal of this research work was aimed to combining the set of the base modified nucleoside (support **7**) with a ribose modification (scheme 4.2).



**Scheme 4.2:** i: **2** or **3** (1.5 eq) in pyridine (1.5 mL/250 mg of resin), DMAP (0.05eq.), 24 h r.t; ii: R-NH<sub>2</sub>/DMF (1:1,w/w), 8 h 50 °C; iii: TFA 5% solution in DCM; iv: EDA/DMF (1:1, w/w) 8 h 50 °C; v: NaIO<sub>4</sub> (50 eq.) in DMF/H<sub>2</sub>O (1:1,v/v), 12 h, 60 °C; resin washings and treatment with NaBH<sub>4</sub> (20 eq.) in EtOH, 2 h, r.t.

As indicative example, the well known C2'-C3' bond oxidative cleavage of the ribose moiety by reaction with metaperiodate followed by the reduction of the di-aldehyde derivative which leads to 2',3'-seconucleosides<sup>[32]</sup>, was examined. In a typical reaction the support **7a-e** (100 mg, 0.13 mmol) was left in contact with a solution of NaIO<sub>4</sub> (1.3 mmol) in DMF/H<sub>2</sub>O (1.5 mL, 1:1,v/v) and shaken for 12 h at 60 °C. The resulting support, after washings with DMF and EtOH, was treated with NaBH<sub>4</sub> (2.6 mmol) in 1.5 mL of EtOH and shaken for 2.0 h at r.t. After washings, the resin **10a-e** was dried under reduced pressure and analyzed by detachment of the nucleoside material by TFA treatment. HPLC analyses indicated that the 2',3'-secoinosine derivatives **11a-e** were obtained in 75-85% yield. The structures of **11a-e** were confirmed by <sup>1</sup>H-NMR (Table 4.1) and MS analyses. The above

oxidative cleavage performed in EtOH/H<sub>2</sub>O (at several ratios and temperatures) afforded the seconucleosides in lower reaction yields most probably due to the scanty swelling of the polystyrene matrix in these solvents mixtures.

#### 4.10. Conclusions

In conclusion, the synthesis of new N-1-dinitrophenyl-inosine based solid supports (**4** and **5**) in which the nucleoside are anchored to a MMT-polystyrene resin by the 5' position, was performed. The supports **4** and **5**, reacting at C-2 position of the purine base with R-NH<sub>2</sub> nucleophiles, were converted in the N-1 alkylated inosine supports (**6** and **7**) and AICAR derivatives supports (**12** and **13**) in very high yields. The detachment of the nucleosidic material, from the above supports, furnished small libraries of N-1 alkylated inosine (**8a-e** and **9a-e**) and AICAR derivatives (**14-15**), having the ribose moiety both protected and unprotected at the 2',3'-hydroxyl functions, in high purities (90-98%). In a further solid-phase reaction, the set of the N-1 alkylated inosine of the supports **7a-e** were combined with the cleavage of the 2',3'-ribose bond. These reactions furnished the new group of solid supports **10a-e** bearing base modified acyclo-nucleosides. Supports **10a-e** released, under acidic conditions, N-1 alkylated-2',3'-secoinosine derivatives **11a-e** in good purity (75-85 %). It is reasonable to suppose that the supports **6**, **7**, **10**, **12** and **13**, bearing a nucleoside derivative can be fruitfully utilized in a combinatorial manner to introduce a number of further derivatizations/conjugations both on the 1-(hydroxyalkyl) function and/or on the ribose or seco-ribose moieties. Briefly, this research work reports the parallel solid-phase combinatorial approach for the rapid synthesis of new N1-alkylated inosine and N1-alkylated-2',3'-secoinosine derivative libraries. MMT-Cl polystyrene resin was confirmed to be an efficient solid support for nucleoside library synthesis. Hydroxyl groups at the 5'-position were preferentially involved into attach the nucleoside scaffold onto the MMT-Cl polystyrene resin. The DNP group is an excellent leaving group for various nucleophilic substitutions on the inosine base to synthesize high quality nucleoside libraries. The method seems good also to prepare a N1-alkylated-2',3'-secoinosine collection. All these new inosine derivatives, with various modifications at N1-position and at 2',3'-bond of the sugar moiety, will be screened against a wide range of biological assays. Further studies are currently in progress in this direction to obtain new and largest libraries of nucleoside analogues.

#### 4.11. Experimental Session

##### 4.11.1 General Methods

NMR spectra were recorded at 400 MHz, and the chemical shifts were expressed relative to the added trimethylsilane. Libraries were analyzed on a HPLC-UV system Jasco PU2089PLUS/UV2075PLUS. A C18 column from Thermo was used for compound purification. The mass spectra at *m/z* 100-1000 were acquired using electrospray ionization with both positive and negative ion detections, by API2000 (Applied Biosystem). UV spectra were recorded at 220-400 nm by a Jasco V530, and the compound purity was monitored based on the UV absorbency at 246-248nm, 267-270nm, 249-250nm, respectively for N1-alkylated-inosine, AICAR derivatives, and N1-alkylated-2',3'-secoinosine.

MMT-Cl polystyrene resin was purchased from Novabiochem. Other starting materials, building blocks, and reagents were purchased from Aldrich and other companies and used directly. The following abbreviations were used throughout the text: DMF = N,N-dimethylformamide, Py = pyridine, MeOH = methanol, EtOH = ethanol, DNP = 2,4-

dinitrophenyl,  $K_2CO_3$  = potassium carbonate, MMT-Cl= 4-methoxytrityl chloride polystyrene resin, DMAP= 4,4'-dimethylaminopyridine,  $NaIO_4$  = sodium m-periodate,  $NaBH_4$  = Sodium borohydride, s's = singlets, d's = doublets, m's = multiplets, q's = quartets

#### 4.11.2. General Procedures

**N1-(2,4-dinitrophenyl)-2',3'-O-isopropylideneinosine (compound 2).** A mixture of **1** (1 g, 2.6 mmol), 2,4-dinitrochlorobenzene (1.87g, 8.1 mmol) and  $K_2CO_3$  ( 897 mg, 6.5 mmol) was stirred in anhydrous DMF (16 mL) at 80 °C for 2.5 h. After cooling, the mixture was filtered and the solid was washed with  $CHCl_3$ . The filtrate and washings, evaporated to dryness under reduced pressure, were purified on a silica gel column eluted with increasing amount of  $CH_3OH$  in  $CHCl_3$  (from 0 to 5%) to give 1.208 g of **2** (as 1: 1 mixture of atropoisomers at *N*-1-phenyl bond) as pale yellow amorphous solid (98% yield).

$^1H$ -NMR (400 MHz) ppm ( $CDCl_3$ ) 1.30 (s, 3H,  $CH_3$ ), 1.45 (s, 3H,  $CH_3$ ), 3.65-4.0 (m, 2H, 2\*5'-H), 4.5 (m, 1H, 4'-H), 5.10 (m, 1H, 3'-H), 5.30 (m, 1H, 2'-H), 6.00 (m, 1H, 1'-H), 7.70 (m, 1H, orto-H), 8.10 (m, 2H, H-8, H-2), 8.70 (m, 1H, meta-H), 9.00 (m, 1H, meta-H); ESI-MS calculated  $m/z$ : 474.11, ( $C_{19}H_{18}N_6O_9$ ), found: 513.50( $M + H$ )<sup>+</sup>.

**N1-(2,4-dinitrophenyl)-inosine (compound 3).** A solution of **2** ( 500 mg, 1.05 mmol) in aqueous  $HCO_2H$  (60%, 14 mL) was stirred at 50°C for 2 h. After the solvent was evaporated under reduced pressure and the residue, dissolved in  $H_2O$  and purified on a C-18 silica gel column eluted with increasing amount of  $CH_3OH$  in  $H_2O$  (from 0 to 10%) to give 451 mg, 99% of **3**.

$^1H$ -NMR (400 MHz) ppm ( $CD_3OD$ ) 3.72-3.95 (m, 2H, 2\*5'-H), 3.9 (m, 1H, 4'-H), 4.15 (m, 1H, 3'-H), 4.35 (m, 1H, 4'-H), 4.60-4.70 (m, 1H, 2'H), 6.10 (m, 1H, 1'-H), 8.00 (m, 1H, orto-H), 8.50 (m, 2H, H-8, H-2), 8.76 (m, 1H, meta-H), 9.06 (m, 1H, meta-H); ESI-MS calculated  $m/z$ : 434.08 ( $C_{16}H_{14}N_6O_9$ ), found: 435,32 ( $M + H$ )<sup>+</sup>.

**Support 4.** A suspension of MMT-Cl resin (500 mg, 0.65 mmol; Novabiochem, loading capacity 1.3 mmol/g), N1-(2,4-dinitrophenyl)-2',3'-O-isopropylideneinosine, **2** (370mg, 0.78mmol, 1.2eq), DMAP (6 mg, 0.05 mmol) in 1.5 mL of anhydrous pyridine was shaken at room temperature for 24h. The resin was filtered and washed sequentially with pyridine-DMF (1:1 v/v, three times) and then with dichloromethane (three times). The thus obtained resin was dried under vacuum for 1h.

**Functionalization.** 1,0 mL of TFA(2% in dichloromethane) was added to 20mg of weight resin. The reaction mixture was shaken for 8 min, then the solution was filtered into a vial. The resin was washed with MeOH, and filtered in the same vial. The solvents were evaporated in vacuo, to evaluate the loading efficiency 85%(1.21mmol/g) based on starting calculated loading capacity of the resin (1.3 mmol/g).

**Support 5.** A suspension of MMT-Cl resin (1g, 1.3 mmol; Novabiochem, loading capacity 1.3 mmol/g), N1-(2,4-dinitrophenyl)-inosine, **3** (676 mg, 1.56 mmol), DMAP (12 mg, 0.1 mmol) in 1.5 mL of anhydrous pyridine was shaken at room temperature for 24h. The resin was filtered and washed sequentially with pyridine-DMF (1:1 v/v, three times) and then with dichloromethane (three times). The thus obtained resin was dried under vacuum for 1h. Loading efficiency 78% (1.01mmol/g).

**Synthesis of N1-alkylated-inosina library**

**Support 6a-f.** Approximately 80 mg (0.1mmol) of starting resin **4** was dispensed in 6 reaction wells using a dispensed spatula and funnel. To each well of resin were added 100 $\mu$ L of anhydrous DMF and 0.5mL (1M) of the (**a-e**) appropriate amine (see Table 1) in DMF. The reaction mixtures were shaken for 4h at 50°C, then filtered and washed with DMF (three times), MeOH-DMF (1:1 v/v, four times), MeOH (three times), MeOH-dichloromethane (1:1 v/v, four times), and finally with dichloromethane (two times). The resultant resins **6a-f** were dried under nitrogen.

**Cleavage. Compounds 8a-e and 14.** To the dried resin in each well was added 2ml of TFA (2% in dichloromethane). The reaction mixtures were shaken for 8 min, then the solution were filtered into a pre-labeled and pre-weighted vials. The resins were washed with MeOH, and filtered in the same corresponding vials. To the combined filtrates was added toluene. The solvents were evaporated under vacuum to provide compounds **8a-e**. The reaction of **4** with ethylenediamine (Table 1, entry f) furnished, as expected, the support **12**, bearing the 2',3'-isopropylidene-AICAR which corresponding detached form **14** was obtained in almost quantitative yields.

**8a=** 98% yields starting from resin **4**;  $^1\text{H-NMR}$  (400 MHz) ppm ( $\text{CD}_3\text{OD}$ ) 0.98 (t, 3H,  $\text{CH}_3$  of  $\text{R}_2$ -moiety), 1.36 (s, 3H,  $\text{CH}_3$ ), 1.6 (s, 3H,  $\text{CH}_3$ ), 1.40-1.76 (m, 4H,  $\text{CH}_2$  of  $\text{R}_2$ -moiety), 3.66-3.80 (m, 2H,  $2^*5'$ -H), 4.12 (t, 2H,  $\text{CH}_2\text{N}$ ), 4.32-4.39 (m, 1H,  $4'$ -H), 5.00 (dd, 1H,  $3'$ -H), 5.26 (dd, 1H,  $2'$ -H), 6.02 (d, 1H,  $1'$ -H), 8.36-8.41 (m, 2H, H-8, H-2). ESI-MS calculated  $m/z$ : 364.17, ( $\text{C}_{17}\text{H}_{24}\text{N}_4\text{O}_5$ ), found: 364.35 ( $\text{M} + \text{H}$ ) $^+$ .

**8b=** 92% yields starting from resin **4**;  $^1\text{H-NMR}$  (400 MHz) ppm ( $\text{CD}_3\text{OD}$ ) 1.36 (s, 3H,  $\text{CH}_3$ ), 1.6 (s, 3H,  $\text{CH}_3$ ), 1.94- 2.04 (m, 2H,  $\text{CH}_2$ ), 3.66-3.80 (m, 2H,  $2^*5'$ -H), 3.82 (t, 2H,  $\text{CH}_2\text{O}$ ), 4.19 (t, 2H,  $\text{CH}_2\text{N}$ ), 4.32-4.39 (m, 1H,  $4'$ -H), 5.00 (dd, 1H,  $3'$ -H), 5.26 (dd, 1H,  $2'$ -H), 6.01 (d, 1H,  $1'$ -H), 8.26-8.41 (m, 2H, H-8, H-2). ESI-MS calculated  $m/z$ : 352.11, ( $\text{C}_{15}\text{H}_{20}\text{N}_4\text{O}_6$ ), found: 353.28, ( $\text{M} + \text{H}$ ) $^+$ .

**8c=** 98% yields starting from resin **4**;  $^1\text{H-NMR}$  (400 MHz) ppm ( $\text{CD}_3\text{OD}$ ) 1.36 (s, 3H,  $\text{CH}_3$ ), 1.6 (s, 3H,  $\text{CH}_3$ ), 1.94- 2.04 (m, 2H,  $\text{CH}_2$ ), 3.60 (t, 2H,  $\text{CH}_2\text{O}$ ), 3.66-3.80 (m, 2H,  $2^*5'$ -H), 4.20 (t, 2H,  $\text{CH}_2\text{N}$ ), 4.32-4.39 (m, 1H,  $4'$ -H), 5.00 (dd, 1H,  $3'$ -H), 5.26 (dd, 1H,  $2'$ -H), 6.02 (d, 1H,  $1'$ -H), 8.32-8.41 (m, 2H, H-8, H-2). ESI-MS calculated  $m/z$ : 366.15, ( $\text{C}_{16}\text{H}_{22}\text{N}_4\text{O}_6$ ), found: 367.32 ( $\text{M} + \text{H}$ ) $^+$ .

**8d=** 98% yields starting from resin **4**;  $^1\text{H-NMR}$  (400 MHz) ppm ( $\text{CD}_3\text{OD}$ ) 1.37 (s, 3H,  $\text{CH}_3$ ), 1.40-1.49 (m, 2H,  $\text{CH}_2$ ), 1.42- 1.81 (m, 9H,  $\text{CH}_3$ ,  $\text{CH}_2$ ), 3.56 (t, 2H,  $\text{CH}_2\text{O}$ ), 3.67-3.80 (m, 2H,  $2^*5'$ -H), 4.12 (t, 2H,  $\text{CH}_2\text{N}$ ), 4.33-4.39 (m, 1H,  $4'$ -H), 5.00 (dd, 1H,  $3'$ -H), 5.26 (dd, 1H,  $2'$ -H), 6.02 (d, 1H,  $1'$ -H), 8.35-8.40 (s, 2H, H-8, H-2). ESI-MS calculated  $m/z$ : 394.19, ( $\text{C}_{18}\text{H}_{26}\text{N}_4\text{O}_6$ ), found: 395.44 ( $\text{M} + \text{H}$ ) $^+$ .

**8e=** 92% yields starting from resin **4**;  $^1\text{H-NMR}$  (400 MHz) ppm ( $\text{CD}_3\text{OD}$ ) 1.37 (s, 3H,  $\text{CH}_3$ ), 1.60 (s, 3H,  $\text{CH}_3$ ), 3.67-3.80 (m, 2H,  $2^*5'$ -H), 3.92 (m, 1H,  $\text{CHN}$ ), 3.98 (d, 4H,  $\text{CH}_2\text{O}$ ), 4.35 (m, 1H,  $4'$ -H), 5.28 (dd, 1H,  $2'$ -H), 6.00 (d, 1H,  $1'$ -H), 8.32-8.38 (s, 2H, H-8, H-2). ESI-MS calculated  $m/z$ : 382.15, ( $\text{C}_{16}\text{H}_{22}\text{N}_4\text{O}_7$ ), found: 382.32 ( $\text{M} + \text{H}$ ) $^+$ .

**14=** 98% yields starting from resin **4**;  $^1\text{H-NMR}$  (400 MHz) ppm ( $\text{CD}_3\text{OD}$ ) 1.37 (s, 3H,  $\text{CH}_3$ ), 1.60 (s, 3H,  $\text{CH}_3$ ), 3.67-3.80 (m, 2H,  $2^*5'$ -H), 4.35 (m, 1H,  $4'$ -H), 5.28 (dd, 1H,  $2'$ -H), 5.66 (d, 1H,  $1'$ -H), 8.04 (s, 1H, H-8). ESI-MS calculated  $m/z$ : 298.13, ( $\text{C}_{12}\text{H}_{18}\text{N}_4\text{O}_5$ ), found: 299.23 ( $\text{M} + \text{H}$ ) $^+$ .

**Support 7a-f.** Approximately 150 mg (0.15mmol) of starting resin **5** was dispensed in 6 reaction wells. To each well of resin were added 100 $\mu$ L of anhydrous DMF and 0.5mL (1M) of the appropriate amine (see Table 1) in DMF. The reaction mixtures were shaken for 4h at 50°C, then filtered and washed with DMF (three times), MeOH-DMF (1:1 v/v, four times), MeOH (three times), MeOH-dichloromethane (1:1 v/v, four times), and finally with dichloromethane (two times). The resultant resins **7a-f** were dried under nitrogen. An half of each one resin (**7a-f**) was dispensed in other 6 reaction wells and cleaved. The



remaining ones were stored as starting material for synthesis of the N1-alkylated-2',3'-secoinosine library.

**Cleavage. Compounds 9a-e and 15.** To the dried resin in each well was added 2ml of TFA (2% in dichloromethane). The reaction mixtures were shaken for 8 min, then the solution were filtered into a pre-labeled and pre-weighted vials. The resins were washed with MeOH, and filtered in the same corresponding vials. To the combined filtrates was added toluene. The solvents were evaporated under vacuum to provide compounds **9a-e**. The reaction of **5** with ethylenediamine (Table 1, entry f) furnished, as expected, the support **13** bearing the AICAR compound, which corresponding detached form **15** was obtained in almost quantitative yields.

**9a=** 98% yields starting from resin **5**;  $^1\text{H-NMR}$  (400 MHz) ppm ( $\text{CD}_3\text{OD}$ ) 0.98 (t, 3H,  $\text{CH}_3$  of  $\text{R}_2$ -moiety), 1.40-1.76 (m, 4H,  $\text{CH}_2$  of  $\text{R}_2$ -moiety), 3.72-3.95 (m, 2H,  $2^*5'$ -H), 4.12 (t, 2H,  $\text{CH}_2\text{N}$ ), 4.15 (m, 1H,  $3'$ -H), 4.35 (m, 1H,  $4'$ -H), 4.60-4.70 (m, 1H,  $2'$ H), 6.02 (m, 1H,  $1'$ -H), 8.36-8.41 (m, 2H, H-8, H-2). ESI-MS calculated  $m/z$ : 324.15, ( $\text{C}_{14}\text{H}_{20}\text{N}_4\text{O}_5$ ), found: 325.35 ( $\text{M} + \text{H}$ ) $^+$ .

**9b=** 95% yields starting from resin **5**;  $^1\text{H-NMR}$  (400 MHz) ppm ( $\text{CD}_3\text{OD}$ ) 3.72-3.95 (m, 2H,  $2^*5'$ -H), 3.82 (t, 2H,  $\text{CH}_2\text{O}$ ), 4.19 (t, 2H,  $\text{CH}_2\text{N}$ ), 4.15 (m, 1H,  $3'$ -H), 4.35 (m, 1H,  $4'$ -H), 4.60-4.70 (m, 1H,  $2'$ H), 6.01 (m, 1H,  $1'$ -H), 8.26-8.42 (m, 2H, H-8, H-2). ESI-MS calculated 312.11, ( $\text{C}_{12}\text{H}_{16}\text{N}_4\text{O}_6$ ), found: 313.28, ( $\text{M} + \text{H}$ ) $^+$ .

**9c=** 96% yields starting from resin **5**;  $^1\text{H-NMR}$  (400 MHz) ppm ( $\text{CD}_3\text{OD}$ ) 1.98 (q, 2H,  $\text{CH}_2$ ), 3.60 (t, 2H,  $\text{CH}_2\text{O}$ ), 3.72-3.95 (m, 2H,  $2^*5'$ -H), 4.20 (t, 2H,  $\text{CH}_2\text{N}$ ), 4.15 (m, 1H,  $3'$ -H), 4.35 (m, 1H,  $4'$ -H), 4.60-4.70 (m, 1H,  $2'$ H), 6.02 (m, 1H,  $1'$ -H), 8.32-8.41 (m, 2H, H-8, H-2). ESI-MS calculated  $m/z$ : 326.12, ( $\text{C}_{13}\text{H}_{18}\text{N}_4\text{O}_6$ ), found: 327.20 ( $\text{M} + \text{H}$ ) $^+$ .

**9d=** 98% yields starting from resin **5**;  $^1\text{H-NMR}$  (400 MHz) ppm ( $\text{CD}_3\text{OD}$ ) 1.42-1.81 (m, 6H,  $\text{CH}_2$ ), 3.56 (t, 2H,  $\text{CH}_2\text{O}$ ), 3.72-3.95 (m, 2H,  $2^*5'$ -H), 4.12 (t, 2H,  $\text{CH}_2\text{N}$ ), 4.15 (m, 1H,  $3'$ -H), 4.35 (m, 1H,  $4'$ -H), 4.60-4.70 (m, 1H,  $2'$ H), 6.02 (m, 1H,  $1'$ -H), 8.35-8.40 (m, 2H, H-8, H-2). ESI-MS calculated  $m/z$ : 354.15, ( $\text{C}_{15}\text{H}_{22}\text{N}_4\text{O}_6$ ), found: 355.45, ( $\text{M} + \text{H}$ ) $^+$ .

**9e=** 90% yields starting from resin **5**;  $^1\text{H-NMR}$  (400 MHz) ppm ( $\text{CD}_3\text{OD}$ ) 3.92 (m, 1H,  $\text{CHN}$ ), 3.98 (d, 4H,  $\text{CH}_2\text{O}$ ), 3.72-3.95 (m, 2H,  $2^*5'$ -H), 4.15 (m, 1H,  $3'$ -H), 4.35 (m, 1H,  $4'$ -H), 4.60-4.70 (m, 1H,  $2'$ H), 6.00 (m, 1H,  $1'$ -H), 8.32-8.38 (m, 2H, H-8, H-2). ESI-MS calculated  $m/z$ : 342.12, ( $\text{C}_{13}\text{H}_{18}\text{N}_4\text{O}_7$ ), found: 342.30, ( $\text{M} + \text{H}$ ) $^+$ .

**15=** 98% yields starting from resin **5**;  $^1\text{H-NMR}$  (400 MHz) ppm ( $\text{CD}_3\text{OD}$ ) 3.72-3.95 (m, 2H,  $2^*5'$ -H), 4.15 (m, 1H,  $3'$ -H), 4.35 (m, 1H,  $4'$ -H), 4.60-4.70 (m, 1H,  $2'$ H), 5.66 (m, 1H,  $1'$ -H), 8.04 (m, 2H, H-8). ESI-MS calculated  $m/z$ : 258.10, ( $\text{C}_9\text{H}_{14}\text{N}_4\text{O}_5$ ), found: 259.30 ( $\text{M} + \text{H}$ ) $^+$ .

### **Synthesis of N1-alkylated-2',3'-secoinosine derivative library**

**Support 10a-e.** The resins **7a-e** (75mg of each one, 0.075mmol) were treated with a solution of  $\text{NaIO}_4$  (168mg, 0.75mmol) in DMF/ $\text{H}_2\text{O}$  (1.5 mL, 1:1 v/v), and shaken for 12h at  $60^\circ\text{C}$ . The resulting supports after washings with DMF and EtOH, were left in contact with a solution of  $\text{NaBH}_4$  (60mg, 1.5 eq) in 1.5mL of EtOH/ $\text{H}_2\text{O}$  for 2h at room temperature. After washings with EtOH, the obtained supports **10a-e** were dried under reduced pressure and analyzed detaching the nucleoside material by TFA treatment.

**Cleavage. Compounds 11a-e.** To the dried resin in each well was added 2ml of TFA (2% in dichloromethane). The reaction mixtures were shaken for 8 min, then the solution were filtered into a pre-labeled and pre-weighted vials. The resins were washed with MeOH, and filtered in the same corresponding vials. To the combined filtrates was added toluene. The solvents were evaporated under vacuum to provide compounds **11a-e**.

**11a=** 85% yields starting from resin **7a**;  $^1\text{H-NMR}$  (400 MHz) ppm ( $\text{CD}_3\text{OD}$ ) 0.96 (t, 3H,  $\text{CH}_3$ ), 1.38-1.75 (m, 4H,  $\text{CH}_2$ ), 3.20 (m, 1H,  $4'$ -H), 3.90 (m, 4H,  $2^*5'$ -H,  $2^*3'$ -H), 4.00 (t, 2H,  $\text{CH}_2\text{N}$ ),

4.22 (d, 2H, 2'H), 6.04 (t, 1H, 1'-H), 8.26-8.30 (m, 2H, H-8, H-2). ESI-MS calculated m/z: 326.16, (C<sub>14</sub>H<sub>22</sub>N<sub>4</sub>O<sub>5</sub>), found: 349.05 (M + Na)<sup>+</sup>.

**11b**= 85% yields starting from resin **7a**; <sup>1</sup>H-NMR (400 MHz) ppm (CD<sub>3</sub>OD) 3.20 (m, 1H, 4'-H), 3.83 (t, 2H, CH<sub>2</sub>O), 3.90 (m, 4H, 2\*5'-H, 2\*3'-H), 4.20 (t, 2H, CH<sub>2</sub>N), 4.40 (d, 2H, 2'H), 6.04 (t, 1H, 1'-H), 8.24-8.28 (m, 2H, H-8, H-2). ESI-MS calculated m/z: 326.16, (C<sub>14</sub>H<sub>22</sub>N<sub>4</sub>O<sub>5</sub>), found: 349.05 (M + Na)<sup>+</sup>.

**11c**= 84% yields starting from resin **7a**; <sup>1</sup>H-NMR (400 MHz) ppm (CD<sub>3</sub>OD) 1.98 (m, 2H, CH<sub>2</sub>), 3.20 (m, 1H, 4'-H), 3.60 (t, 2H, CH<sub>2</sub>O), 3.90 (m, 4H, 2\*5'-H, 2\*3'-H), 4.21 (t, 2H, CH<sub>2</sub>N), 4.40 (d, 2H, 2'H), 6.03 (t, 1H, 1'-H), 8.26-8.31 (m, 2H, H-8, H-2). ESI-MS calculated m/z: 326.16, (C<sub>14</sub>H<sub>22</sub>N<sub>4</sub>O<sub>5</sub>), found: 349.05 (M + Na)<sup>+</sup>.

**11d**= 82% yields starting from resin **7a**; <sup>1</sup>H-NMR (400 MHz) ppm (CD<sub>3</sub>OD) 1.43-1.79 (m, 6H, CH<sub>2</sub>), 3.20 (m, 1H, 4'-H), 3.55 (t, 2H, CH<sub>2</sub>O), 3.90 (m, 4H, 2\*5'-H, 2\*3'-H), 4.11 (t, 2H, CH<sub>2</sub>N), 4.40 (d, 2H, 2'H), 6.03 (t, 1H, 1'-H), 8.25-8.28 (m, 2H, H-8, H-2). ESI-MS calculated m/z: 326.16, (C<sub>14</sub>H<sub>22</sub>N<sub>4</sub>O<sub>5</sub>), found: 349.05 (M + Na)<sup>+</sup>.

**11e**= 75% yields starting from resin **7a**; <sup>1</sup>H-NMR (400 MHz) ppm (CD<sub>3</sub>OD) 3.20 (m, 1H, 4'-H), 3.90 (m, 4H, 2\*5'-H, 2\*3'-H), 3.95 (q, 1H, CHN), 4.02 (d, 4H, CH<sub>2</sub>O), 4.40 (d, 2H, 2'H), 6.05 (t, 1H, 1'-H), 8.29-8.04 (m, 2H, H-8, H-2). ESI-MS calculated m/z: 326.16, (C<sub>14</sub>H<sub>22</sub>N<sub>4</sub>O<sub>5</sub>), found: 349.05 (M + Na)<sup>+</sup>.

## References

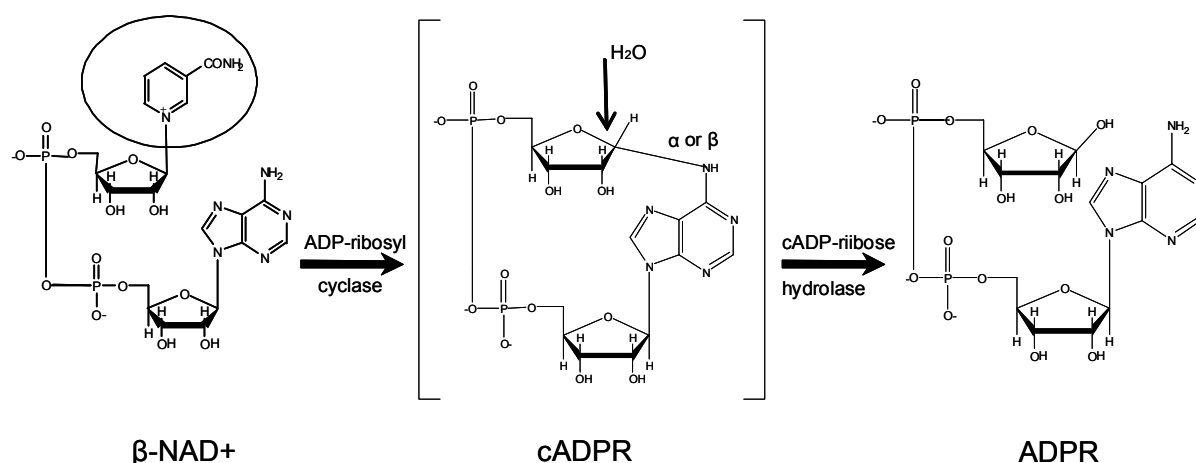
1. (a) Miura, S.; Izuta, S. *Current Drug Targets* **2004**, 5, 191-195; (b) Parker, W.B.; Secrist, J.A.; Waud, W.R. *Current Opinion in Investigational Drugs*, **2004**, 5, 592-596; (c) Szafraniec, S.I.; Stachnick, K.J.; Skierski, J.S. *Acta Poloniae Pharmaceutica* **2004**, 61, 297-306.
2. (a) Lagoja, I.M. *Chemistry & Biodiversity*, **2005**, 2, 1-50; (b) Kimura K.; Bugg, T.D.H. *Nat. Prod. Rep.* **2003**, 20, 252-273; (c) Rachakonda, S.; Cartee, L. *Current Medicinal Chemistry*, **2004**, 11, 775-793.
3. (a) In *Antiviral Nucleosides*; Chu, C.K., Ed.; Elsevier: , 2003; (b) Simons, C.; Wu, Q.; Htar, T.T. *Current Topics in Medicinal Chemistry* **2005**, 5, 1191-1203; (c) De Clercq, E.; Neyts, J. *Rev. Med. Virol.* **2004**, 14, 289-300.
4. Mathè, C.; Gosselin, G.; *Antiviral research* , **2006**, 71, 276-281
5. Pastor-Anglada, M.; Cano-Soldado, P.; *Virus Research*, **2005**, 107, 151-164
6. Schinazi, R.; Hernandez-Santiago, B.; *Antiviral research* , **2006**, 71, 322-334
7. Zoulim, F.; *Journal of Antimicrobial Chemotherapy*, **2005**, 55, 608-611
8. Plunkett, W., Huang, P.; *Semin. Oncol.*, **1996**, 22 (4 Suppl 11), 3-10
9. Plunkett, W., Huang, P.; *Semin. Oncol.*, **1996**, 23 (5 Suppl 10), 3-15
10. Rivkin, A.M.; *Ann.Pharmacother.*, **2004**, 38, 625-633
11. Houghten, R.A.; Pinilla, C.; *J.Med.Chem.*, **1999**, 42, 3743
12. An, H.; Cook, P.D.; *Chem.Rev.*, **2000**, 100, 3311
13. Reed, N.N.; Janda, K.D.; *Org.Lett.*, **2000**, 2, 1311
14. Nakamura, H.; Linclau, B.; *J.Am.Chem.Soc.*, **2001**, 123, 10119
15. Jin, Y.; Chen, X.L.; *Bioorg.Med.Chem.Lett.*, **2001**, 11, 2057
16. Becauge, S.L.; *Methods Mol. Biol.*, **1993**, 20, 33
17. Ding, Y.; Habib, Q.; Shaw, S.Z.; Li, D.Y.; . *J. Comb. Chem.* **2003**, 5, 851-859.
18. Rodenko, B.; Wanner, M.J.; *J. Chem. Soc., Perkin Trans. 1* **2002**, 1247-1252
19. Epple, R.; Kudirka, R.; Greenberg, W.A. *J. Comb. Chem.* **2003**, 5, 292-310.
20. de Champdoré M.; De Napoli, L.; et al.; *Chem. Commun.* **1997**, 2079-2082.
21. De Napoli, L.; Di Fabio, G.; et al.; *D. Synlett* **2004**, 11, 1975-1979.
22. Zhou, W.; Roland, A.; et al.; *Tetrahedron Lett.* **2000**, 41, 441-445.
23. Jin, Y.; Roland, A.; et al.; *Bioorganic & Medicinal Chemistry Letters* **2000**, 10, 1921-1925
24. Roland, A.; Xiao, Y.; Jin, Y.; Iyer, R.P. *Tetrahedron Lett.* **2001**, 42, 3669-3672
25. Jin, Y.; Chen, X., et al.; *Bioorganic & Medicinal Chemistry Letters* **2001**, 11, 2057-2060.
26. De Napoli, L.; Messere, A.; Montesarchio, D.; Piccialli, G. *J. Org. Chem.* **1995**, 60, 2251-2253.
27. De Napoli, L.; Messere, A.; et al.; *J. Chem. Soc., Perkin Trans. 1* **1997**, 2079-2082.
28. N Ariza, X.; Bou, V.; Villarasa, J. *J. Am. Chem. Soc.* **1995**, 117, 3665-3673
29. Terraza, M.; Ariza, X.; Farràs, J.; Villarrasa, J.; *Chem. Commun.* **2005**, 3968-3970
30. Narukulla, R.; Shuker, D.E.G.; Xu, Y.-Z. *Nucleic Acids Research* **2005**, 33, 1767-1778
31. Galeone, A.; Mayol, L., et al.; *Eur. J. Org. Chem.* **2002**, 4234-4238
32. Kumar, A.; Walker, R.T.; *Tetrahedron*, **1990**, 3101-3110

## Chapter 5

***“Synthesis of a new series of cADPR (cyclic hypoxanthine-diphospho-ribose) analogues, designed as novel stable mimics of cADPR a potent  $\text{Ca}^{2+}$ -mobilizing second messenger”***

### Introduction

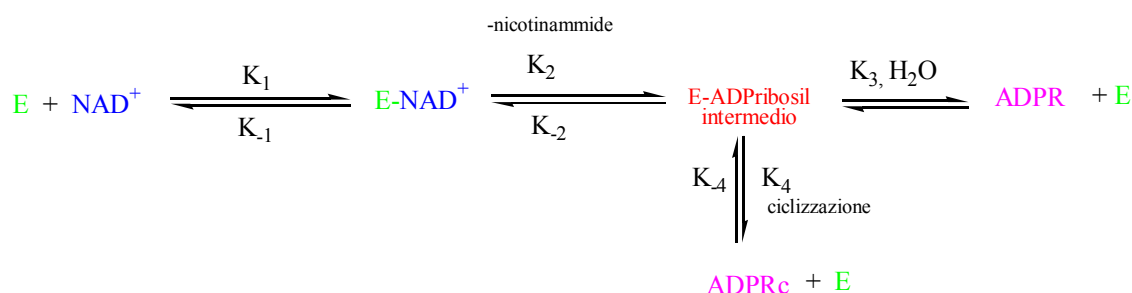
Cyclic ADP-ribose (cADPR) was discovered in 1987 when Lee and co-workers<sup>[1]</sup> examined various metabolites for their abilities to induce calcium release in sea urchin egg homogenates. Observing pyridine nucleotide levels fluctuated markedly after fertilization, they tested  $\text{NAD}^+$  released almost as much calcium as  $\text{IP}_3$ , but there was a time lag of 1-4 min. Since  $\text{NAD}^+$  itself does not induce calcium release, the delayed response was later shown to reflect the time for its enzymatic conversion to a novel metabolite, which was isolated by HPLC in quantities sufficient for structural determination. Using NMR and mass spectroscopy, Lee and co-workers<sup>[1]</sup> suggested that this metabolite is a cyclic compound with the adenine ring of the  $\text{NAD}^+$  molecule forming a N-glycosidic linkage between the N6-amino group of the adenine with the anomeric carbon of the terminal ribosyl unit by displacing the nicotinamide group (figure 5.1).



**Figure 5.1:** Lee's first proposed structure of cADPR.

cADPR has the characteristic 18-membered cycling structure consisting of an adenine, two (N1- and N9-) riboses, and a pyrophosphate, in which the two primary hydroxyl groups of the riboses are linked by a pyrophosphate unit. Under neutral conditions, cADPR is in a zwitterionic form with a positive charge around the N(1)-C(6)-N6 moiety ( $\text{pK}_a=8.3$ ), making the molecule extremely unstable. The charged adenine moiety attached to the anomeric carbon of the N1-ribose can be a very efficient leaving group. Since the metabolite is a cyclic compound and can be hydrolyzed to ADP-ribose (ADPR), the name cyclic-ADP-ribose was assigned to this compound with the proposed gross structure as shown in figure 5.1. The absolute stereochemistry of the linkage between the adenine and the terminal ribose ( $\alpha$  or  $\beta$ -linkage) was left unassigned. Cyclic ADP-ribose has been shown to mobilize intracellular  $\text{Ca}^{2+}$  in various cells, such as sea urchin eggs, pancreatic  $\beta$ -cells, smooth muscles, cardiac muscles, T-lymphocytes, and cerebellar neurons, indicating that it is a general mediator involved in  $\text{Ca}^{2+}$  signalling<sup>[2]</sup>. Indeed, ADP-ribosyl-

cyclase activities are present in many mammalian and invertebrate tissues, suggesting that cADPR may function widely as a calcium mobilizing agent. Since cADPR is a newly identified metabolite, it is appropriate to group all the cADPR synthesizing enzymes together in a new family of ADP-ribosyl cyclases. ADP-ribosyl cyclase cyclizes NAD to produce cyclic ADP-ribose (cADPR) and releases nicotinamide in the process. It is an ubiquitous enzyme widely present in animal cells, and has been identified in *Euglena*, a protozoan. The first cyclase purified was from *Aplysia ovotestis*. It is a soluble protein of approximately 30 KDa. Sequence comparison reveals that CD38, a lymphocyte antigen, is a mammalian homolog. CD38 is a transmembrane glycoprotein of approximately 45 KDa. Additionally, both CD38 and the *Aplysia* cyclase use NADP as a substrate and catalyze the exchange of the nicotinamide group with nicotinic acid, producing nicotinic acid adenine dinucleotide phosphate (NAADP). In addition to the *Aplysia* cyclase and CD38, CD157 (BST1), another antigen, is also a homolog. The three proteins share 25-30% sequence identity and there is also a perfect alignment of the cysteines of the homologs. However, CD38 and CD157 have very similar catalytic properties, possessing both cADPR synthesizing and hydrolyzing activities, while the other two members, the *Aplysia* and *Euglena* cyclases, lack the hydrolase activity. The dominant enzymatic product of CD38 and CD157 from NAD is thus ADP-ribose, with cADPR being only a minor component (figure 5.2).



**Figure 5.2:** Mechanism of CD38 activity.

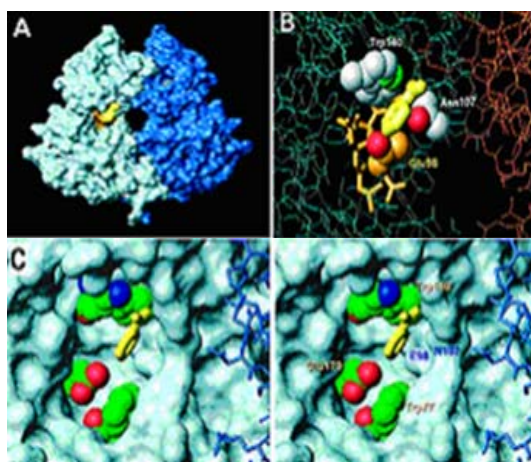
The overall reaction is similar to classical NADases. *Aplysia* and *Euglena* cyclases, are clearly different from NADase as they produce cADPR instead of ADP-ribose. However, ADP-ribosyl cyclase can be easily distinguished from other NADases and the cyclases can be operationally classified. In contrast to cyclase, cADPR hydrolases are also widely distributed along mammalian tissues, where they hydrolyze cADPR at the N-1 ribosyl linkage to give the inactive ADPR. Biological as well as chemical cADPR instability (discussed below), limits studies of its physiological role. Therefore, stable analogues of cADPR that have similar  $\text{Ca}^{+2}$ -mobilizing activity to that cADPR are required.

### 5.1. *Cyclic-ADP-Ribose: a new way to control calcium*

Cyclic adenosine diphosphate ribose (cADPR) is a novel  $\text{Ca}^{2+}$ -releasing second messenger candidate. As a natural occurring metabolite of nicotinamide adenine dinucleotide ( $\text{NAD}^+$ ), it mobilizes calcium by a mechanism completely independent of the  $\text{IP}_3$  receptor<sup>[3]</sup>. When a second class of intracellular calcium release channels, the ryanodine receptor, was found in sea urchin eggs<sup>[4]</sup>, the possibility that cADPR was an endogenous regulator of ryanodine receptors arose. Ryanodine receptors were first characterized as the major calcium release channel of muscle sarcoplasmic reticulum and share substantial similarities with  $\text{IP}_3$  receptors in that they are both transmembrane proteins with large cytoplasmic domains and a quatrefoil arrangement of four identical subunits<sup>[5]</sup>. A key property of ryanodine receptors, first recognized from studies on isolated sarcoplasmic reticulum and then confirmed by channel isolation and reconstitution in artificial bilayers, is that they can be opened by calcium, that is, they can mediate calcium-induced calcium-release (CICR)<sup>[5]</sup>. Ryanodine receptors are widespread in many cell types and have been incorporated in many models to explain the spatial and temporal complexity of cellular calcium signals, which can be oscillatory and can propagate as regenerative planar or spiral waves<sup>[6]</sup>. On the basis of cross-desensitization studies and experiments with known ryanodine receptor antagonist (ryanodine, ruthenium red and procaine), cADPR was shown to release calcium via this ryanodine-sensitive CICR mechanism in the sea urchin eggs, suggesting that cADPR is an endogenous regulator of CICR through a ryanodine receptor rather than through an  $\text{IP}_3$  receptor<sup>[4]</sup>. This hypothesis has been supported by the demonstration that cADPR at nanomolar concentrations, but not  $\text{IP}_3$ , mimics caffeine (a ryanodine receptor activator) in inducing oscillations in a calcium dependent ion current known to reflect intracellular calcium oscillations in rat dorsal root ganglion cells<sup>[7]</sup>. Although cADPR has been shown to have calcium mobilizing activity in a number of cell types, including sea urchin eggs, pituitary cells and dorsal root ganglion cells, it has not been clear whether cADPR functions as a second messenger for extracellular stimuli. Rather, cADPR could be a factor for CICR, simply enhancing the sensitivity of its receptor to calcium; this notion would be consistent with the fact that calcium and cADPR potentiate the effects of each other and that cADPR itself is present in many mammalian cell types under resting conditions. However, the major advance in identifying cADPR as a new calcium-mobilizing second messenger has come from the finding that glucose induces a rise in cADPR concentrations in pancreatic  $\beta$ -cells, and moreover cADPR stimulates insulin secretion from permeabilized  $\beta$ -cells<sup>[8]</sup>. Furthermore, cADPR seems to release calcium from  $\beta$ -cell microsomes via a ryanodine-sensitive mechanism, consistent with the proposed mechanism for this agent's actions in regulating CICR. Although many agents have been suggested as important intracellular mediators of nutrient-induced insulin secretion, cADPR is particularly attractive because a role for this molecule in stimulus-secretion coupling in the  $\beta$ -cell can explain the mechanisms of a class of diabetogenic drugs, the  $\beta$ -cytotoxins such as streptozotocin. These agents deplete intracellular  $\text{NAD}^+$ , the precursor of cADPR.

## 5.2. ADP-ribosil cyclase

ADP-ribosil cyclase is a dimer, in which each monomer was a bean-shaped molecule consisting of two domains separated by a central cleft<sup>[9]</sup>. The two monomers associated with each other in a head-to-head fashion forming a donout-like molecule with a central cavity, as shown by chemical cross-linking and dynamic laser scattering. The tentative assignment of the bound nicotinamide is shown in yellow in the figure 5.3A. The binding site was in a pocket at the central cleft of the monomer and was very close to the central cavity of the dimer. Nicotinamide was bound to each monomer. The binding pocket of the monomer colored dark blue was not visible, because it was on the other side of the monomer. The cyclase has a stretch of six amino acids (TLEDTL) that are highly conserved with CD38. This conserved region formed part of the binding pocket and is shown in orange in figure 5.3A. Figure 5.3B shows the three amino acids in the binding pocket that were closest to the placement of the bound nicotinamide (yellow) and appeared to complex the bound substrate.



**Figure 5.3:** Molecular modelling of site active of ADPribosylcyclase.

These three residues were Trp140 (2.8Å°), Glu98 (3.5Å°) and Asn107 (3.1Å°). The conserved sequence is shown as orange sticks, and Glu 98 is part of this sequence. Trp140, although not in the conserved region, is nevertheless conserved among two species of the cyclase and three species of CD38. Asn107, on the other hand, is not a conserved residue. If these amino acids are indeed responsible for substrate binding, their alteration by site-direct mutagenesis should result in reduction or elimination of the enzymatic activities of the cyclase. Data indicates that the enzyme activity has an exquisitely specific requirement for both the charge and size of the chain at the 98 position, and any alteration would result in decrease in activity. Of the three amino acids that complex the nicotinamide, Trp140 is the most critical to the cyclase activity. It is closest to the bound nicotinamide, and its conversion to glycine resulted in a 4,648-fold decrease in activity. Asn107 is the least critical; only a 2-fold decrease in activity was measured for the mutant N107G. Asp99 is next to Glu98 and is similarly charged. It is also in the highly conserved sequence, TLEDTL. Nevertheless its conversion to glycine produced much less pronounced effect on the cyclase activity, only a 10-fold decrease. These results indicate that the effect of the site-direct mutagenesis is residue specific and that Trp140 is the most critical residue for the cyclase activity. To search for the catalytic residues, other amino acids that make-up the binding pocket were mutagenized. The

results indicate that Glu179 and also Trp77 are most likely the catalytic residues. The critical importance of Glu179 was further shown by substituting it with various other residues. All substitutions resulted in inactivation of the enzyme. X-ray crystallography shows that both Trp77 and Trp140 line the rim of the binding pocket, one on each side. Particularly, Trp140 is in close proximity to the bound nicotinamide, suggesting that it is important in positioning the substrate, while Trp77 may serve a similar function in positioning the substrate  $\text{NAD}^+$  through hydrophobic interaction. Another important function that the two tryptophan residues may serve is to curtail the access of water to the catalytic site through the hydrophobicity of their side chains. Inaccessibility of water to the active site would ensure that the nucleophilic attack of the oxocarbenium intermediate by the 1-nitrogen of the adenine is the dominant reaction, leading to the cyclization of  $\text{NAD}^+$  to produce cADPR. Access of water to the active site, on other hand, would lead to the formation of ADP-ribose through hydrolysis, which is the dominant reaction catalyzed by CD38. Figure 5.3C shows a stereo view of the active site. The van der Waals contact surface of the site is shown with the three most critical amino acids, Trp140, Trp77, and Glu179, rendered as space filling models. The active site is in a pocket with Glu179 deep inside the pocket, one on each side. The carboxyl oxygen atoms of Glu179 and the carbonyl oxygen atom of Trp77, which are likely to participate directly in catalysis, are shown in red. The bound nicotinamide is shown in yellow. The aromatic ring of Trp140 is close to, and appears to position, the bound nicotinamide, presumably through hydrophobic interaction. Another notable structural feature of the active site is its close proximity to the central cavity formed by the two monomer cyclase molecules. The size of the central cavity is similar to the dimension of a molecule of cADPR. This suggests the possibility that the central cavity may serve as a channel, through which the product of catalysis, cADPR, can pass.

### 5.3. CD38

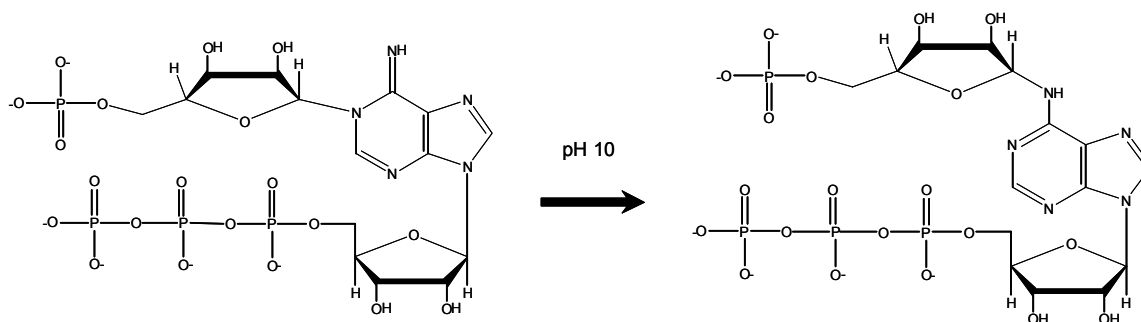
The crystal structure of CD38 has not been determined. However, it is possible to model the structure of its extra-membrane domain using the crystal coordinates of the *Aplysia* cyclase because of the sequence homology between the proteins. The resulting energy-minimized structure reveals that the two extra pairs of cysteines, Cys119 and Cys201 are close enough to form a disulfide bond<sup>[10]</sup>. More importantly, it also reveals that the residues in CD38, Glu226, Trp125 and Trp189, corresponding to the critical amino acids of the cyclase are all clustered in a pocket near the middle of the molecule, very similar to the active site pocket of *Aplysia* cyclase. This is remarkable considering that these residues spread over a third of the sequence of the molecule. Mutagenizing Glu226 to even highly conserved residues, such as aspartate, results in complete loss of all enzymatic activities<sup>[11]</sup>. Another critical residue is Glu146<sup>[12]</sup>. It is part of the conserved sequence. Glu146 is located directly across from the catalytic residue, Glu 226. This is not situated as deep inside the active site pocket as Glu 226 and is closer to the position of the coordinating tryptophane, Trp189, which is at the rim of the pocket. The distance between Glu226 and Glu146 is about  $8\text{\AA}$ . Homology modelling is used to assess the structural changes at the active site caused by the mutations. Replacement of Glu146 with alanine greatly increases the cyclase activity and depresses the hydrolase activity. The side chain of alanine is smaller than glutamate. The distance between Ala146 and Glu226 is thus larger, at  $9.8\text{\AA}$ , making the opening of the active site pocket slightly wider. The net effect of these changes likely allows the substrate to bind deeper into the pocket. Glu146 substitution with amino acids that increase hydrophobicity, greatly enhances cyclase activity. For example, substituting Glu146 of CD38 with phenylalanine (E146F) enhances



the cyclase activity as much as the E146A mutant. The main difference between the two mutants is that E146F has higher hydrolase activity than E146A. The size of the side chain of the phenylalanine is larger than glutamine and the distance between Phe146 and Glu226 is now closer at 7.5 Å. The substitution is likely to reduce access to the catalytic residue<sup>[13]</sup>. CD38 have also provided strong experimental support that it may function as such a novel transporter of cADPR. It is shown that when resealed right-side-out red cell ghosts that contain ecto-CD38 are incubated with NAD, active transport of cADPR occurs, generating a concentration gradient of the product 10-80 fold higher inside than outside<sup>[14]</sup>. Inhibition of the enzymatic activity of cADPR by mercaptoethanol blocks the process. Similar active transport of the enzymatic products is seen with purified CD38 reconstituted into liposomes. That CD38 is a transporter capable of active translocation of its enzymatic products across membranes provides a novel explanation of how a surface antigen with its catalytic domain on the outside of the cells can participate in intracellular  $\text{Ca}^{2+}$  signalling.

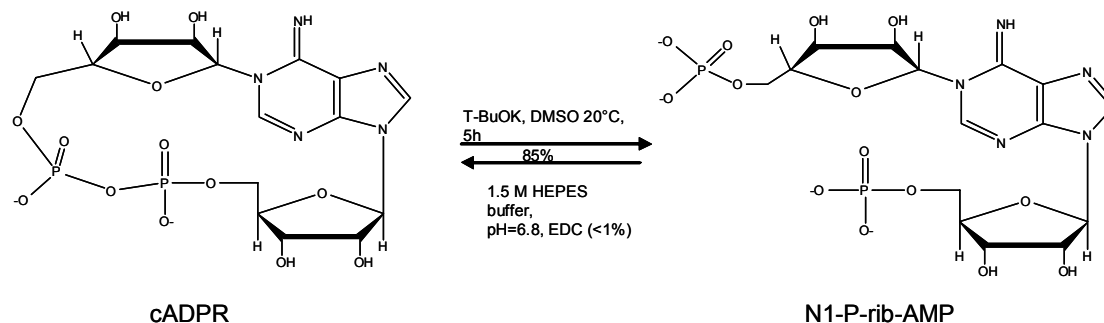
#### 5.4. Stability of cADPR

cADPR is characterized by a very labile N1-glycosidic bond, which is rapidly hydrolyzed both enzymatically, by cADP-hydrolase<sup>[15]</sup>, than non-enzymatically, to give ADP-ribose even in a neutral aqueous solution<sup>[16]</sup>. It is slowly hydrolyzed to ADPR at room temperature under slightly acidic conditions with a half-life of about 10 days. At 37°C, the half-life for spontaneous hydrolysis is about 24h. The pH titration curve of cADPR revealed that it has a pK<sub>a</sub> of 8.2 and  $\epsilon=3200$  at 290 nm. The absence of a pK<sub>a</sub> at pH=4.0, corresponding to the pK<sub>a</sub> of the amino group of ADPR or AMP, indicates that cADPR, like N1-P-rib-ATP forms an imino base rather than an amino base<sup>[17]</sup>. It is well documented that N1-alkylated adenosine derivatives undergo the Dimroth rearrangement in base to give the N6-derivatives<sup>[18]</sup>, (figure 5.4).



**Figure 5.4:** Base-catalyzed transformation of N1-P-rib-ATP to N6-P-rib-ATP.

Numerous mechanistic studies, including  $^{15}\text{N}$ -isotope labelling, have shown that this rearrangement proceeds through adenine ring opening without cleavage of the N1-alkyl bond. Indeed Ames<sup>[19]</sup> has observed that N1-P-rib-ATP, an open ring analogue of cADPR, underwent such a rearrangement at pH 10. On the other hand, it is interesting to note that cADPR is very stable in base and does not undergo the Dimroth rearrangement at pH 10. The pyrophosphate bond in cADPR is also moderately resistant to alkaline hydrolysis, but can be cleaved to N1-P-rib-AMP with potassium t-butoxide in dimethyl sulfoxide<sup>[20]</sup> (figure 5.5).



**Figure 5.5:** Base –catalyzed cleavage of cADPR.

The pyrophosphate linkage in cADPR is resistant to the action of pyrophosphatases in contrast to ADPR, which is readily attacked by these enzymes<sup>[21]</sup>. Biological as well as chemical instability of cADPR limits studies of its physiological role. Therefore, stable analogues of cADPR that have a  $\text{Ca}^{2+}$  mobilizing activity in cells similar to that of cADPR are urgently required.

### 5.5. Synthesis of cADPR analogues

cADPR analogues can be used in investigating the mechanism of cADPR-mediated  $\text{Ca}^{2+}$  signalling pathways. Agonists or antagonists of cADPR are also expected to be lead structures for the development of drugs, since cADPR has been shown to play important physiological roles, such as insulin release from  $\beta$ -cells<sup>[8]</sup>. Therefore the synthesis of cADPR analogues has been extensively investigated. At least, cADPR analogues have been synthesized predominantly by enzymatic and chemoenzymatic methods using ADP-ribosyl-cyclase-catalyzed cyclization under mild neutral conditions. By this method, cADPR analogues modified in the purine ring and the N9-ribose moieties have been prepared from the corresponding  $\beta$ -nicotinamide-di-nucleotide-type precursors<sup>[22]</sup>. Although evaluation of these analogues has identified biologically useful ones, the analogues obtained by these methods are limited due to the substrate specificity of the ADP-ribosyl-cyclase. Then, previous studies of enzymatically and chemoenzymatically synthesized cADPR analogues have disclosed the structure-activity relationship (SAR) for the N9-ribose and adenine moieties<sup>[23]</sup>, at least to some extent. Nevertheless, the SAR of the N1-ribose moiety remained virtually unknown, since N1-ribose-modified analogues are difficult to access by the chemoenzymatic and enzymatic route. In the enzymatic and chemoenzymatic synthesis, preparation of sugar-modified NAD-type precursors is crucial. Furthermore, even if these sugar-modified NAD analogues are provided, the cyclase may still not catalyze their desired cyclization. Therefore, a chemical synthetic method is essential for providing these N1-ribose-modified analogues to clarify the SAR and also to develop compounds of further biological importance. In recent years, on the other hand, methods for the chemical synthesis of cADPR analogues have been extensively studied, and useful cADPR analogues, which could not be prepared by the enzymatic and chemoenzymatic methods, have been synthesized. In the synthesis of cADPR and its analogues, construction of the large 18-membered ring structure is the key step, and different methods have been developed.

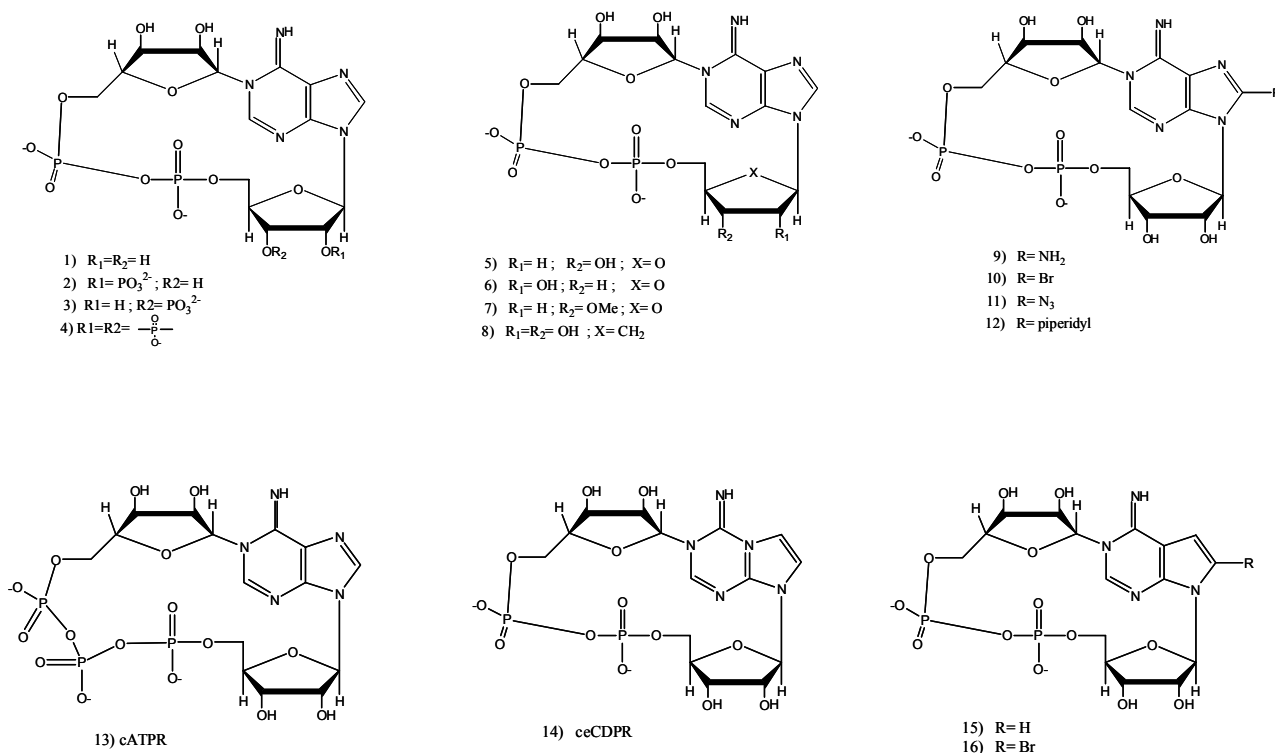
### 5.5.1. *Enzymatic synthesis of cADPR derivatives*

There are three types of enzyme systems known for the conversion of  $\text{NAD}^+$  into cADPR (cyclase) or ADPR (NADase) and the hydrolysis of cADPR to ADPR (hydrolase):

1. A membrane-bound  $\text{NAD}^+$  glycohydrolase (NADase) was purified to homogeneity from canine spleen microsomes, which contained three enzyme activities. The ratio of NADase:cyclase:hydrolase activities for this enzyme was found to be 100:2:30<sup>[24]</sup>.
2. The human and rodent CD38 antigen also possess cyclase and hydrolase activities. An ectoenzyme was purified to homogeneity from solubilized human erythrocyte membranes and exhibited a similar ratio of enzyme activities (100:1:10)<sup>[25]</sup>.
3. In contrast, the enzyme derived from the invertebrate, *Aplysia Californica* is very rich in cyclase activity and has little NADase and hydrolase activities<sup>[26]</sup>.

This enzyme, namely ADP-ribosyl cyclase, and is now commercially available from Sigma in a highly purified state, and has been widely used for the laboratory preparation of cADPR and its analogues. The crystal structure of the cyclase was reported in 1996<sup>[27]</sup>. This enzyme has been also cloned in the methylotrophic yeast, *Pichia pastoris* that expressed high levels of this enzyme<sup>[28]</sup>. Using high biomass fermentation up to 300 milligram per liter of cyclase was achieved. SDS-PAGE analysis revealed that the heterologous protein comprised 90-95% of the total protein secreted extracellularly. The properties of the recombinant enzyme compared favourably to those of the native enzyme. The yeast expression system can produce gram quantities of this cyclase. Hence the supply of the enzyme for synthetic applications is no longer a problem. The ADP-ribosyl cyclase of *Aplysia Californica* converts  $\beta\text{-NAD}^+$  analogues into three different types of products depending on the structural features of substrates. In all cases,  $\alpha\text{-NAD}^+$  derivatives are not cyclized or hydrolyzed by this enzyme (figure 5.6).

**Cyclization at N-1:** The ADP-ribosyl cyclase efficiently converts  $\beta\text{-NAD}^+$  into cADPR(1) in over 90% yield because this enzyme has little NADase and hydrolase activities. In a similar manner, it transforms  $\text{NADP}^+$  into the corresponding cADPRP(2), 3'- $\text{NADP}^+$  into 3'-cADPRP(3) and 2',3' cyclic  $\text{NADP}^+$  into 2',3'-cADPRP(4)<sup>[29]</sup>. cADPRP stimulates calcium release two times more effectively than cADPR, but by a mechanism distinct from both cADPR and  $\text{IP}_3$ . The other two analogues are inactive in inducing calcium release. The occurrence of cADPRP in mammalian tissues was also demonstrated<sup>[29]</sup>. Due to the wide distribution of  $\text{NADP}^+$  in mammalian systems and its unique calcium releasing property, several papers on the role of cADPRP in the regulation of  $\text{Ca}^{2+}$  homeostasis and its characterization have been reported<sup>[30]</sup>. The 2'-deoxy, 3'-deoxy and 3'-O-methyl derivatives of  $\text{NAD}^+$  were converted by cyclase into their respective cADPR derivatives<sup>[31]</sup>. 2'-deoxy-cADPR(5) is similar to cADPR in inducing  $\text{Ca}^{2+}$  release, whereas 3'-deoxy-cADPR(6) is 100-fold less potent. However, 3'-O-methyl-cADPR(7) inhibited the cADPR-induced  $\text{Ca}^{2+}$  release in a dose-dependent manner. This suggested that the 3' hydroxy-, but not the 2'-hydroxy group was essential for the calcium releasing activity. Cyclic aristeromycin diphosphate ribose (cArisDPR, 8), the first carbocyclic analogue of cADPR, was prepared via the cyclization of  $\text{NArisD}^+$ <sup>[32]</sup>. Interestingly, while retaining a similar calcium release profile to that of cADPR, cArisDPR showed higher stability towards enzymatic hydrolysis. All the above results indicated that the cyclase can tolerate substantial changes around the ribosyl moiety of the adenosine portion of the  $\text{NAD}^+$  derivatives. As the cyclase enzyme can accommodate large substituents at the 8-position of the adenine ring of  $\text{NAD}^+$  analogues, a series of 8-substituted cADPR analogues were prepared. Among these, 8-amino-cADPR(9), was the strongest antagonist of cADPR, whereas 8-Br-(10) and 8-azido-cADPR(11) were less potent as antagonists, and 8-piperidyl-cADPR(12) was inactive<sup>[33]</sup>.

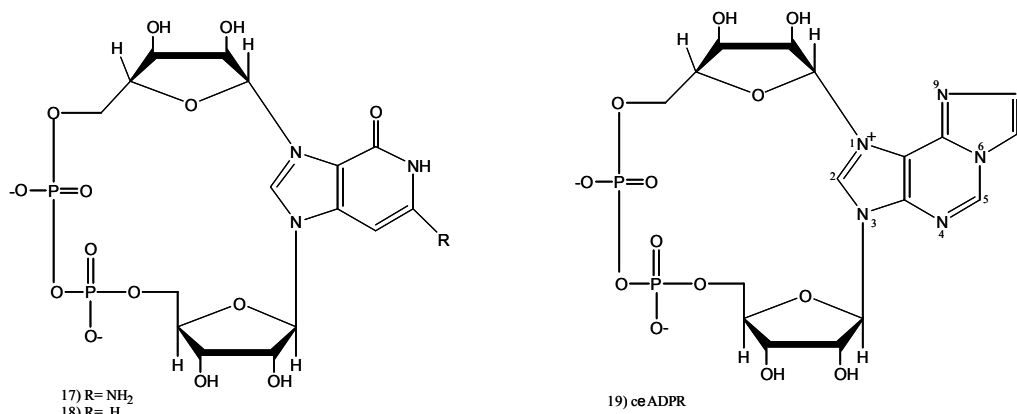


**Figure 5.6:** cADPR analogues cyclized at the N-1 position.

The triphosphate analogues of cADPR (cATPR, **13**) was formed via the cyclization of 5'-triphosphopyridine nucleotide (5'-TPN)<sup>[34]</sup>. cATPR was about 20 times more potent in inducing  $Ca^{2+}$  release and was more resistant than cADPR to hydrolysis by NAD-glycohydrolase. More importantly, cATPR and cADPR modulate the  $Ca^{2+}$  release via the same mechanism. These properties enabled cATPR to be used as a suitable molecular probe for the study of cADPR binding proteins. The cyclization of nicotinamide cytosine dinucleotide using the *Aplysia* cyclase was attempted, also, but no cyclized product was detected, which suggested that a bicyclic ring system may be required for cyclization. This assumption led to deployment of 3,4-ethenocytosine ring system as a mimic of the adenine ring of  $NAD^+$ . Indeed, this nicotinamide-3,4-ethenocytosine diphosphate analogues was converted into a new cyclic nucleotide, cyclic etheno-CDP-ribose (cε-CDPR, **14**) in 92% yield<sup>[35]</sup>. This result poignantly reflected the importance of substrate structure in dictating product formation. Although cε-CDPR was found to be inactive in inducing  $Ca^{2+}$  release, it is the first cADPR analogue with a modified purine ring. Comparing the structural features of the 3,4-ethenocytosine ring versus that of the adenine ring, it is worth noting that the cyclization site of cε-CDPR (N-1) corresponded to the N-9 position of adenine ring. This means the direction of cyclization of  $NAD^+$  analogues could occur not only from N-9 to N-1 but also possibly from N-1 to N-9 for the generation of unique cADPR analogues. cADPR analogues modified at N-7, such as 7-deaza-cADPR(**15**) and 7-deaza-8-bromo-cADPR(**16**)<sup>[36]</sup> were also prepared via cyclization of the corresponding  $NAD^+$  analogues. The former was a partial agonist whereas the latter was an antagonist of cADPR. However, both were resistant to chemical and enzymatic hydrolyses and possessed the important feature of membrane permeability.

**Cyclization at N-7:** The cyclase of *Aplysia Californica* catalyzed an alternative mode of cyclization of nicotinamide guanine dinucleotide ( $NGD^+$ ) and nicotinamide hypoxanthine

dinucleotide ( $\text{NHD}^+$ ) to form cyclic GDP-ribose (cGDPR, **17**) and cyclic HDP-ribose (cHDPR, **18**), respectively (figure 5.7)<sup>[37]</sup>.

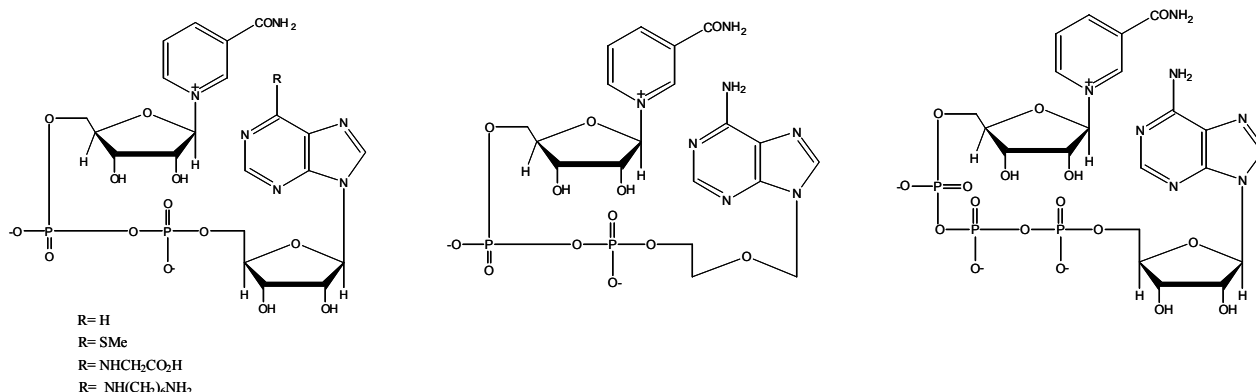


**Figure 5.7:** cADPR analogues cyclized at N-7.

In these cyclic nucleotides, the newly formed glycosyl bonds were shown to be attached onto the N-7 nitrogen of the purine rings and not to the N-1 position of the guanine ring as previously proposed<sup>[38]</sup>. The cyclase also catalyzed the transformation of nicotinamide - 1,N6-etheno-adenine dinucleotide into a novel cyclic nucleotide (cε-ADPR, **19**) whose N-glycosil bond was attached onto the N-1 position of the 1,N6-etheno-adenine nucleus corresponding to the N-7 position of the adenine ring<sup>[39]</sup>. While all of the three N-7 cyclized cADPR analogues were much less active than cADPR in inducing calcium release in the rat microsomal system, their novel ring structures were more stable than cADPR towards heat and NADase hydrolysis. Furthermore, their unique fluorescent properties provide investigators with a useful tool for monitoring the cyclase reaction especially in crude homogenates<sup>[40]</sup>. The cyclization of  $\text{NGD}^+$  to cGDPR may be used as a continuous spectrofluorimetric assay of cyclase activity. This assay method has several advantages.  $\text{NGD}^+$  is cyclized to cGDPR in yield higher than that of  $\text{NAD}^+$  to cADPR, the fluorescent chromophore is formed only upon cyclization offering low background and high sensitivity, the assay is continuous and is adaptable to spectrophotometric or fluorimetric detection.

### 5.5.2. *Non-enzymatic syntheses of cADPR analogues*

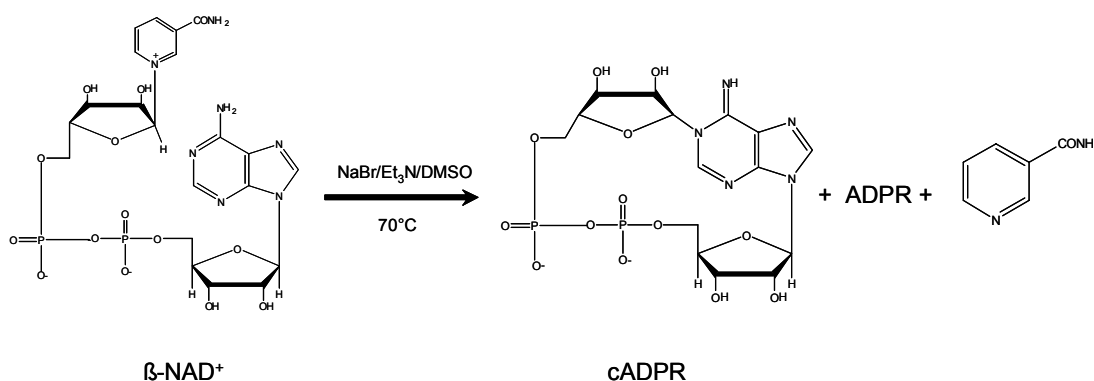
Since the discovery of cADPR, much attention has been paid to its synthesis by non-enzymatic methods. The goal of the non-enzymatic methods was not only to verify the structure of cADPR, but also to prepare stable cADPR analogues or analogues that were not possible to prepare by the enzymatic method (figure 5.8).



**Figure 5.8:** NAD<sup>+</sup> analogues not cyclized by ADP-ribosyl cyclase

Until now, two strategies have been developed for the synthesis of cADPR and its analogues:

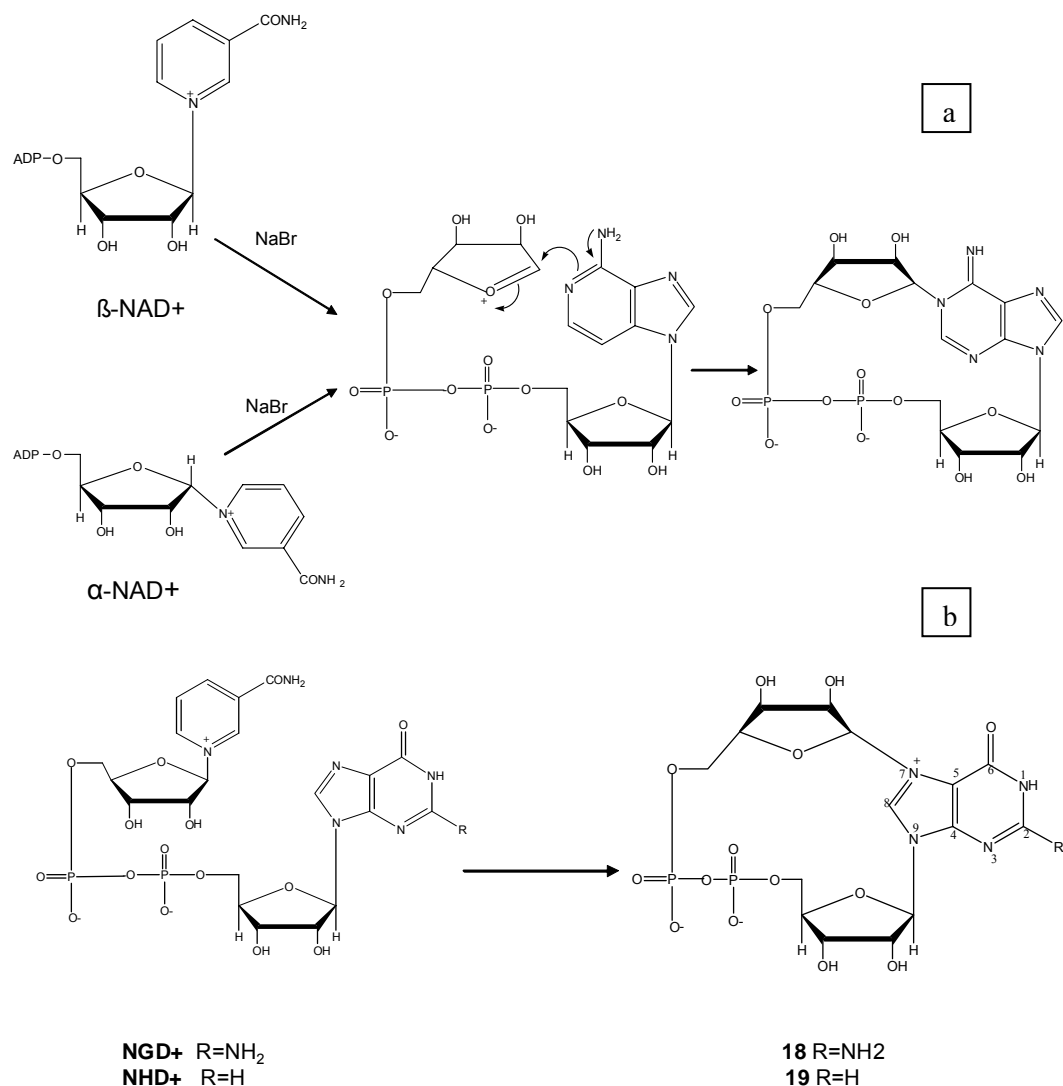
**Biomimetic synthesis of cADPR.** Since de novo synthesis of cADPR would require extensive chemoselective protection and deprotection of functional groups, Yamada et al.<sup>[41]</sup> developed a biomimetic cyclization of  $\alpha$ - and  $\beta$ -NAD<sup>+</sup>. It was discovered that by reacting  $\beta$ -NAD<sup>+</sup> with NaBr and triethylamine in dry DMSO at 70°C, cADPR was obtained in 28% yield (figure 5.9).



**Figure 5.9:** Biomimetic cyclization of  $\beta$ -NAD<sup>+</sup>.

Lower yields were obtained using other halide salts. It is noteworthy that this intramolecular cyclization was completely stereoselective as cADPR was obtained as the sole isomer and it proceeded via a common oxo-carbenium ion intermediate (figure 5.10a). To determine the versatility of this cyclization reaction, a series of NAD<sup>+</sup> analogues were treated under the same reaction conditions. In most cases biomimetic route gives similar product profiles to that of enzymatic method, but higher yields are generally obtained by the latter method. 2'-deoxy-cADPR, c $\epsilon$ -CDPR and cATPR were obtained via N-1-cyclization, whereas cGDPR(**17**) and cHDPR(**18**) were obtained via N-7 cyclization (figure 5.10b). However, one exception was when nicotinamide- N6-etheno-adenine dinucleotide ( $\epsilon$ -NAD<sup>+</sup>) was subjected to the condition of biomimetic synthesis, no formation of c $\epsilon$ -ADPR was detected. Instead, a new cyclic nucleotide (9- c $\epsilon$ -ADPR) was obtained<sup>[39]</sup> and its structure was assigned as a cyclic product in which the newly formed glycosyl bond was attached to N-9 instead of N-1 position. It was about two times more potent than cADPR in inducing Ca<sup>+2</sup> release and its unique fluorescence behaviour, provides investigators with a

useful probe for the study of cADPR-binding proteins. This is the first time that it is possible to observe that the enzymatic and biomimetic methods proceeded via different reaction pathways to give different cyclization products, thereby, demonstrating the usefulness of the biomimetic method.

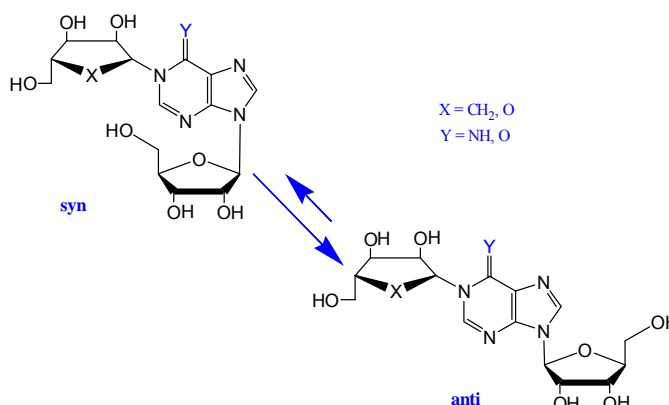


**Figure 5.10:** Mechanism of biomimetic cyclization of  $\alpha$  and  $\beta$ -NAD<sup>+</sup> via a) N-1cyclization; b) N-7cyclization.

### Total synthesis of cADPR analogues via intramolecular condensation reaction

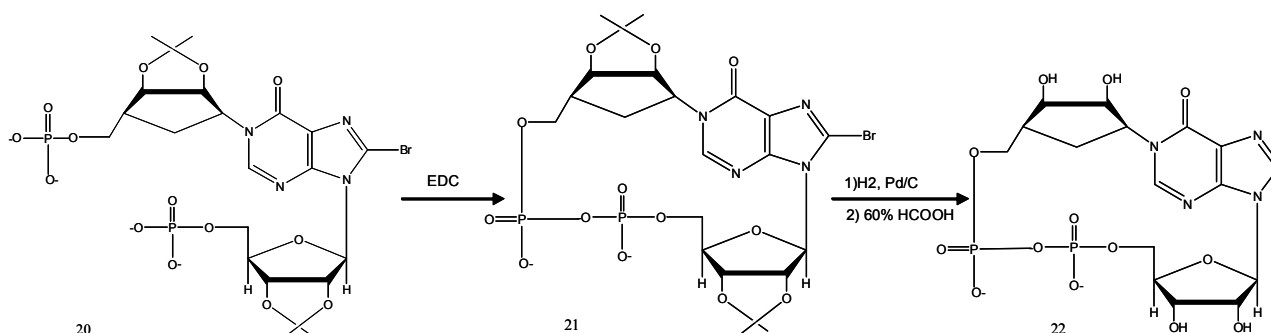
This method was first used by Zhang and coworkers to prepare cADPR in order to confirm its structure. The key intermediate N1-P-rib-AMP was prepared by a chemoenzymatic method. Its cyclization to cADPR in the presence of EDC was attempted but was not successful, due to the low cyclization yield (<1%). Later, several hydrolysis-resistant carbocyclic analogues of N1-P-rib-AMP were prepared. The ring closure was attempted but failed. Later successful cyclization of analogues of N1-P-rib-AMP, using a modified method of Shuto et al.<sup>[42]</sup>, was achieved. It was surmised that the intramolecular cyclization would depend on the conformations of the nucleotides. Nucleotides could exist as the *anti* and *syn* conformers (figure 5.11). In the *syn* conformer, the two phosphate groups are close to each other, which should facilitate the intramolecular condensation reaction. Unfortunately, it is well known that the *anti* conformation predominates in most natural nucleotides, which may be the reason for the unsuccessful cyclization of N1-P-rib-AMP.

By introducing a bulky group at the 8-position of purine nucleosides, the conformation of nucleotides may be restricted to the **syn** form due to the steric interaction between the bulky group and the ribose moiety.



**Figure 5.11:** Possible nucleoside conformations.

Based on this assumption, Shuto et al.<sup>[43]</sup> synthesized compound (**20**), an 8-bromo substituted analogue of N1-P-rib-IMP, which was cyclized to afford (**21**). After removal of the 8-bromo substituent and the protecting groups, the cADPR analogue (**22**) was obtained (figure 5.12). This constituted the first cADPR analogue synthesized by a total synthetic approach, and provided a general method for the preparation of stable cADPR analogues. However, the overall yield was low (23%).

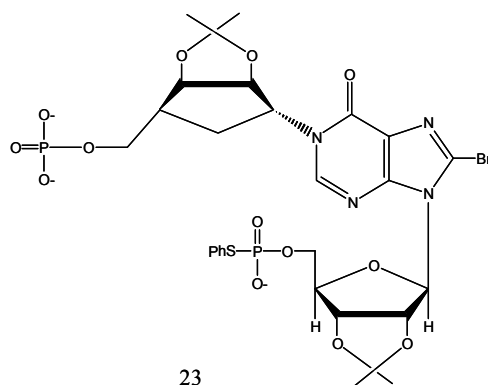


**Figure 5.12:** The first total synthesis of a cADPR analogue.

## 5.6. Chemical cyclization strategies for new cADPR analogues synthesis

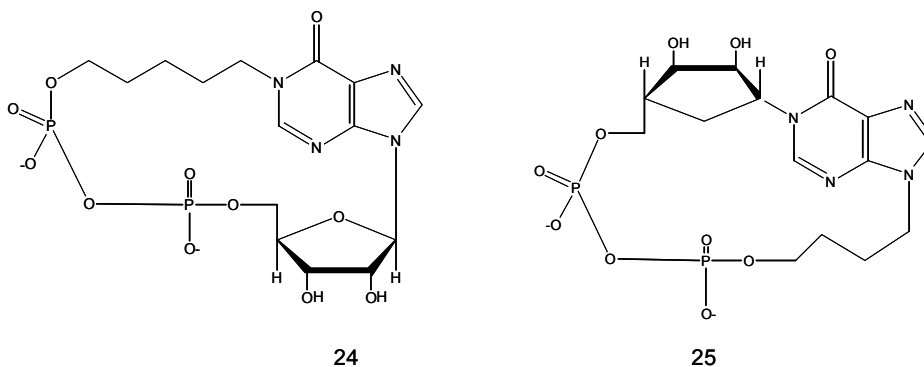
Chemical synthesis of cADPR and its analogues also show limitation due to the regio- and stereo-selective formation of N1-glycosidic bond and to the intramolecular cyclization involving the pyrophosphate bond realization. The difficulty of forming such an intramolecular pyrophosphate linkage, which prevented the completion of the synthesis of target cADPR analogues, has also been experienced of other groups. Thus, the development of an efficient condensation method for forming the intramolecular pyrophosphate linkage was needed. Matsuda and co-workers proposed a new synthetic plan to obtain cIDPcR (**22**), in which formation of the intramolecular pyrophosphate linkage was completed using  $\text{AgNO}_3$  or  $\text{I}_2/3\text{MS}$  as promoter and 5'-phenylthiophosphate-N1-P-carbaribose-inosine (**23**, figure 5.13) as substrate for cyclizing<sup>[44]</sup>.





**Figure 5.13:** Phenylthiophosphate-type substrate.

Adopting Matsuda's condensation procedure, in Piccialli's laboratories has been synthesized a new cIDPR analogue (**24**, figure 5.14), characterized by the presence of carbaribosyl and butyl moieties at N-1 and N-9 of hypoxanthine, respectively<sup>[45]</sup>.



**Figure 5.14:** N1, N9-ribose modified cIDPR.

This compound was the first example of a new class of molecules characterized by a linear chain at N-9, designed to investigate the role of the N-9 attached ribosyl moiety in the mechanism of  $\text{Ca}^{+2}$  intracellular modulation. Then, the synthesis of a new cIDPR analogue (**25**) having a penthyl chain at N-1 of an inosine unit, has been realized<sup>[46]</sup>. In this latter case an alternative intramolecular condensation reaction was performed to cyclize **a** and **b** (figure 5.15), respectively the emiprotected and unprotected linear bisphosphate precursors of **25**. In the former method **a**, dissolved in pyridine, was added over 15h to  $\text{I}_2$ , dissolved in the same solvent, in the presence of molecular sieves ( $3\text{\AA}$ ), to furnish the cyclic product **25** (60% isolated yield). In the latter, the intramolecular condensation reaction of the two phosphate groups of **b** was performed by addition of a slight excess of EDC in MPD and allowing the mixture to stand at  $50^\circ\text{C}$  for 60 h. The desired cyclic product **25** was obtained in 80% yield after HPLC purification. The surprisingly and unprecedented high cyclization yields observed using EDC as a condensing agent of the two phosphomonoester functions of the linear precursor **b**, indicates this reaction as an efficient and easy method to obtain large rings characterized by a high conformational mobility. Previous computational studies followed using SOFTWARE INSIGHT II 2000, and as forcing drive the CVFF, minimized with the conjugated alghorithm, RMS derivative 0.001; DINAMIC: EQUILIBRATION 4000 STEPS, STEPS 4000, History 100, at  $323^\circ\text{K}$  and using MPD as solvent with a dielectric constant of 32, suggested that the N1-penthyl cIDPR derivative adopts a conformation in which the phosphates distance is  $12.10\text{ \AA}$ , a

smallest value if we consider cIDPR. This latter strategy will considerably facilitate the preparation of further cADPR analogues as it is both fast and economical.

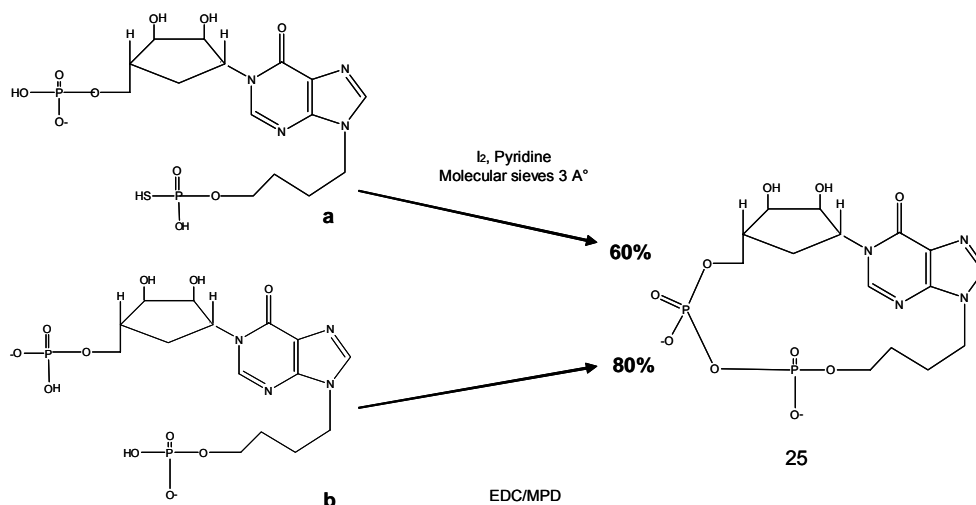


Figure 5.15: Intramolecular condensation strategies.

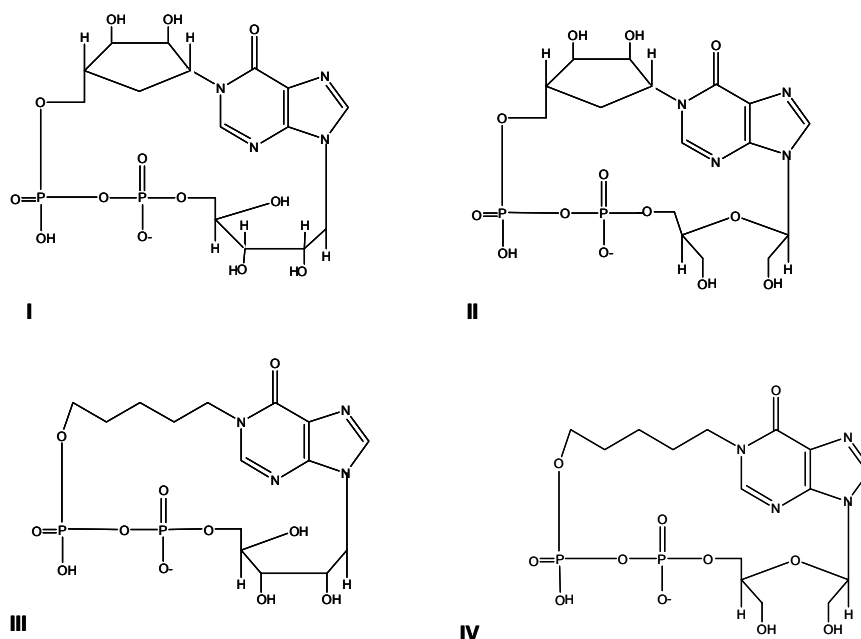
### 5.7. Structural cADPR analogues for SAR elucidation

Intensive studies of cADPR have been made and needed because of its biological importance not only as pharmacological tool, but also in the elucidation of SAR, of mechanism of cADPR action and labelling of binding proteins. Since cADPR is rapidly hydrolyzed, many non-hydrolyzable mimics of cADPR, which retained a similar calcium release profile to that of cADPR are reported. Numerous structural modifications of cADPR described include the replacement of the adenine portion with a deaza-ring system<sup>[47]</sup>, the modification or deletion of either one of the ribose units, as well as combinations of the above structural features. Additional structural modification of cADPR includes the pyrophosphate moiety. Cyclic adenosine triphosphate ribose was more potent in inducing  $Ca^{2+}$  release as compared to cADPR<sup>[34]</sup>. Since the N1-ribosyl moiety and the 6-amino group in cADPR are not critical structural factors for retaining the biological activity, cyclic IDP-ribose analogues have been synthesized. Shuto and co-workers first described the chemical synthesis of cyclic IDP-carbocyclic-ribose (cIDP-carbaribose)<sup>[43]</sup>. cIDPR analogues were found to retain the cADPR activities to induce  $Ca^{2+}$  release in T-cells<sup>[48]</sup>. N1 and/or N9 glycosyl-substituted cyclic IDP-ribose derivatives were previously reported in our laboratory, the N1-pentyl cyclic inosine diphosphate ribose and the N1-carbocyclic, N9-butyl analogue (**24**, **25**). These compounds were designed to investigate the role of the N-1 and N-9 attached ribosyl moiety in the mechanism of  $Ca^{2+}$  intracellular modulation. Preliminary tests to evaluate their biological effects on SK-N-BE cells of a human neuroblastoma system, have been performed.  $[Ca^{2+}]_i$  releasing activity was observed only when compound **25** the N1-pentyl analogue was added, while compound **24** didn't show activity both in the same and higher concentrations. Other 3'-modified cIDPR derivatives were also synthesized, their biological response suggested that the 3'-OH was essential for the calcium releasing activity.

### 5.8. The aim of the work

To further collect more informations on southern N-9 ribose role (more precisely hydroxyl groups and/or cycle rigidity occurrence) into  $\text{Ca}^{2+}$  release activity of cyclic IDP-ribose derivatives, the synthesis of N-9 acyclic sugar derivatives have been recently performed. Many different bond scissions of the furanosyl endocyclic frame can be chemoselectively accomplished to produce a variety of acyclic derivatives.

The oxidative cleavage of 2', 3'-cis-diol portion of inosine with  $\text{NaIO}_4$ <sup>[49]</sup>, followed by reduction with  $\text{NaBH}_4$ , and the selective cleavage of the C1'-O4' bond in the ribose ring with  $\text{DIBAL-H}$ <sup>[50]</sup> have been performed, to obtain 2',3'-seco- and N9-ribityl-cADPR analogues, respectively (Figure 5.16).



**Figure 5.16:** New synthesized cIDPR derivatives.

I wish to report here the synthesis of new cIDPR analogues (*I*, *II*, *III*, *IV*), which are expected to be resistant to both enzymatic and chemical hydrolysis as are similar N-1-alkylated analogues, and possessing respectively:

- *I* a hydroxypentyl chain at N-9 and a carbaribose at N-1 of the hypoxantine base;
- *II* differs from *I* for the presence of a pentyl chain at N-1 position of the hypoxanthine base;
- *III* a polihydroxylate ether strand at N-9 and a carbaribose at N-1 of the hypoxantine base;
- *IV* differs from *III* at N-1 position of the hypoxanthine base, where the carbaribose moiety is substituted by a pentyl chain.

## 5.9. Chemistry

The new compounds (*I*;*II*;*III*;*IV*) have been synthesized from the inosine base, differently opened at N9-ribose, and alternatively substituted at N1 position with a carbaribose or a penthanol unit. The key steps for the synthesis of these analogues involve ribose opening strategy, substitution at N1-position of inosine and intramolecular cyclization of corresponding bisphosphate, following Piccialli's method. The syntheses of compounds *I*; *II* and *III*; *IV* are summarized in Scheme 5.1 and 5.2, respectively.

### *I* and *II* compounds:

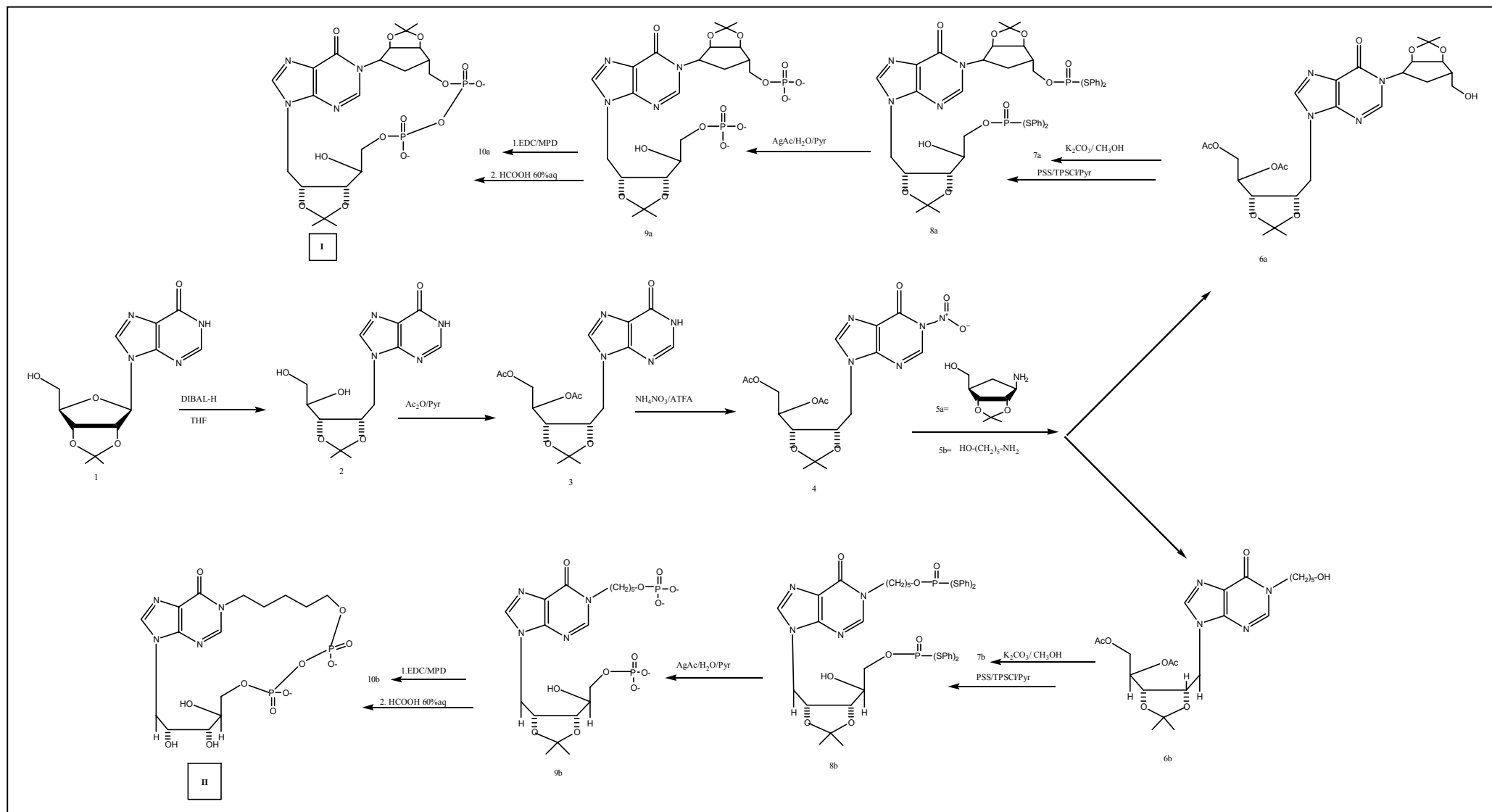
The synthetic strategy adopted for *I* and *II* is shown in Scheme 5.1. The 2',3'-O-isopropylidene-inosine **1** was allowed to react with DIBAL-H in anhydrous THF<sup>[50]</sup>.

In this way the sugar moiety was reduced to afford the corresponding 1-D-ribitylinosine derivative **2**. Acetylation of the hydroxyl functions of **2** gave derivative **3**. N-1 nitroinosine derivative **4** was prepared according to the procedure proposed by Vilarresa and co-workers by treatment of **3** with ammonium nitrate and trifluoroacetic anhydride (TFAA) in dichloromethane<sup>[51]</sup>. Derivatives **6a** and **6b** have been obtained by reaction of **4** with **5a** and **5b**, respectively, in DMF. Removal of the acetyl protecting groups from **6a** and **6b** by treatment with K<sub>2</sub>CO<sub>3</sub> in MeOH, yielded **7a** and **7b**, respectively. The bis(phenylthio)phosphoryl groups were introduced on the primary hydroxyl functions by treatment of **7a** and **7b** with S,S-diphenyl-phosphorodithioate (cyclohexylammonium salt) in the presence of 2,4,6-triisopropylbenzenesulphonylchloride (TPSCI) and tetrazole in pyridine<sup>[52]</sup> to give the protected bisphosphate derivative **8a** and **8b** as the main product together with the triphosphate derivatives in a 8:2 ratio. After HPLC (RP-C18) purification, **8a** and **8b** were treated with silver acetate (AgAc) for complete removal of protecting groups of both phosphate residues<sup>[52]</sup>. The deprotected bisphosphate derivative **9a** and **9b** were further purified by HPLC (RP-C18) before cyclization, which was performed as previously described, to obtain **10a** and **10b** (55 % yield after purification). Finally, these latters were deprotected at 2',3'-hydroxyl groups by treatment with formic acid, which afforded the target compounds *I* and *II*, respectively from **10a** and **10b** (15 % overall yield starting from **1**), whose structures were confirmed by <sup>1</sup>H and MS data.

### *III* and *IV* compounds:

The synthetic strategy adopted to obtain compounds *III* and *IV* is reported in Scheme 5.2. The inosine **1'**, in anhydrous pyridine, was treated with dimethoxytrytilchloride (DMT-Cl) to afford 5'-(O-DMT)-inosine (**2'**), as main product. 1-chloro-2,4-dinitrobenzene and anhydrous K<sub>2</sub>CO<sub>3</sub> were added to a solution of **2'** in dry DMF to afford the N1-(2,4-dinitro)-phenyl derivative **3'**. This intermediate was rapidly transformed in **4'a** or **4'b** using **a** or **b** (10 equiv in DMF, 80 °C), respectively. Then the oxidative cleavage of 2',3'-cis-diol portion of inosine with NaIO<sub>4</sub>, followed by reduction of aldehydic groups with NaBH<sub>4</sub><sup>[49]</sup> was performed and afforded compound **5'a** and **5'b**. The unprotected 2',3'-seco-nucleosides **5'a** and **5'b** were treated with acetic anhydride to lead fully protected compound **6'a** and **6'b** in high yields. These latters were treated with 2% trifluoroacetic acid (TFA) in DCM to remove the dimethoxytrytil groups (**7'a** and **7'b**). The unprotected hydroxyl groups were converted into corresponding bis-(phenylthio)-phosphoryl derivatives (**8'a** and **8'b**) treating **7'a** and **7'b** with S,S-diphenyl-phosphorodithioate (cyclohexylammonium salt) in the presence of 2,4,6-triisopropylbenzenesulphonylchloride (TPSCI in pyridine). After purification, **8'a** and **8'b** were treated with silver acetate (AgAc) for complete removal of protecting groups from both phosphate residues (**9'a** and **9'b**).

## CHAPTER 5 Synthesis of cIDPR analogues



**Scheme 5.1:** Synthetic scheme followed to obtain compounds **I** and **II**

The deprotected bisphosphate derivative **9'a** and **9'b** were further purified by HPLC (RP-C18) before cyclization, which was performed with EDC/NMP, to obtain **10'a** and **10'b** (50 % yield after purification). Finally, these latter were deprotected at hydroxyl groups by treatment with anhydrous  $K_2CO_3$  to remove the 2', 3'-acetyl groups, common at both **10'a** and **10'b**, followed by  $HCOOH$  60%<sub>(aq)</sub> treatment, only for **10'a**. The target compounds **III** and **IV** were obtained respectively from **10'a** and **10'b** (36 % overall yield starting from **1**), and their structures were confirmed by  $^1H$  NMR and MS data.

### 5.10. Conclusions

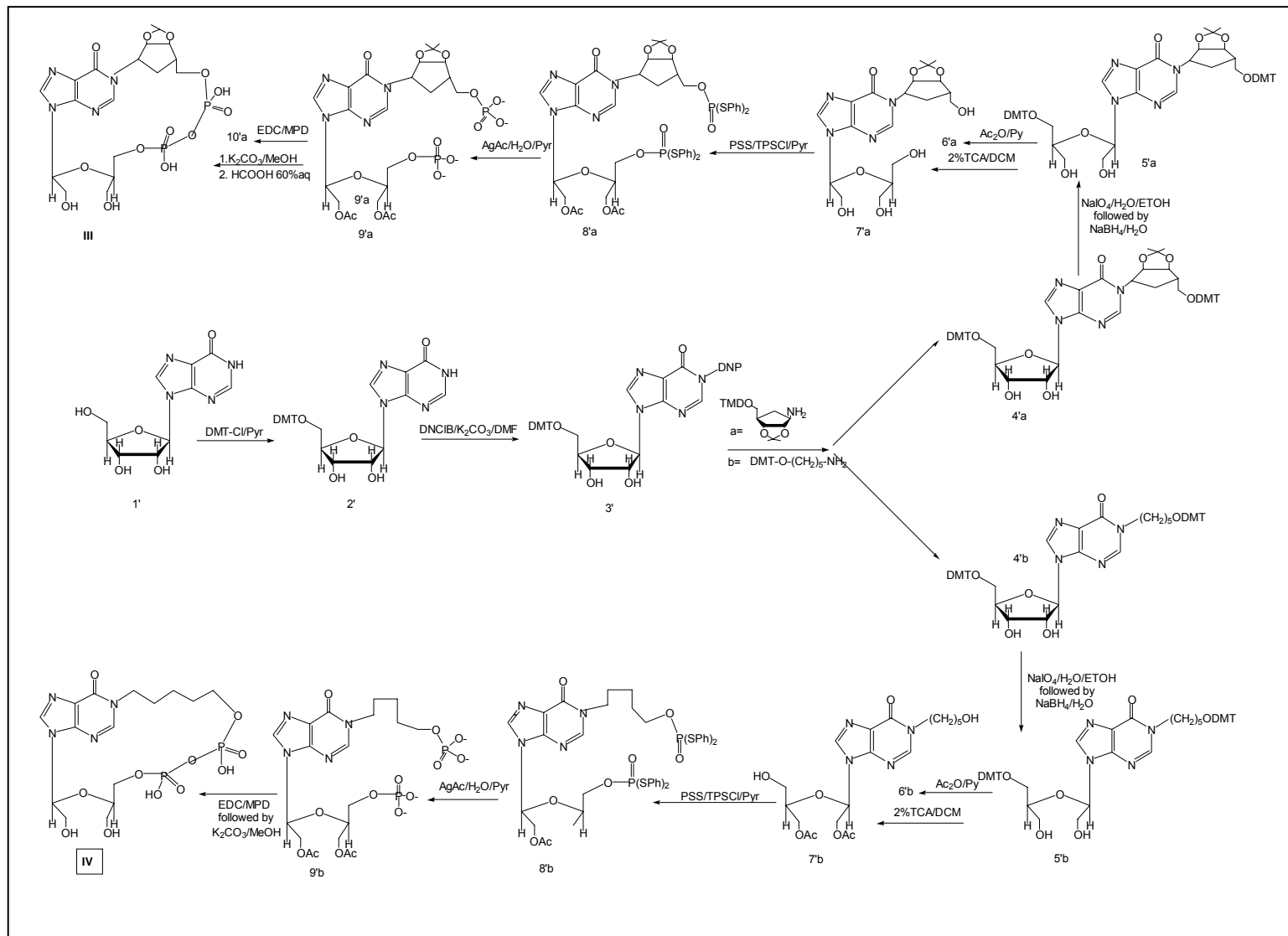
The above compounds were designed to investigate the role that N-1 and N-9 attached ribosyl moieties play in the mechanism of intracellular  $Ca^{2+}$  modulation. The structure of these compounds were confirmed by UV,  $^1H$ -NMR and MS, and they will be tested in their  $Ca^{2+}$ -mobilizing activity. The evaluated mobilizing effect of these newly synthesized compounds would be effective in understanding the SAR and it will provide a greater demonstrating of the southern ribose role in the ligand-receptor recognition, both about the importance of hydroxyl groups presence and ribose cycle rigidity need. These results could lead to novel therapeutic agents.

### 5.11. Experimental Session

#### 5.11.1. General Methods

$^1H$  and  $^{13}C$  NMR spectra were recorded in  $CDCl_3$  on a Bruker WM 500 spectrometer. Residual proton and carbon signals of the solvent ( $CDCl_3$   $\delta$ =7.24 and 77.5;  $CD_3OD$   $\delta$  = 3.30 and 77.5;  $D_2O$   $\delta$  = 4.60) were used as internal references.  $^1H$  NMR signals were assigned to the pertinent nuclei through two-dimensional  $^1H$ - $^1H$  and  $^1H$ - $^{13}C$  COSY experiments. Mass spectra were registered by an API-2000 of Applied Biosystem instruments. General reagents and solvents were purchased from Sigma-Aldrich-Fluka. UV measurements were carried out on a Jasco V-530 UV spectrophotometer. HPLC purifications were performed on a Nucleosil C18 column (Macherey Nagel, 250/10, 7  $\mu m$ ). The following abbreviations were used throughout the text:  $NH_4NO_3$  = ammonium nitrate, TPSCI = 2,4,6-triisopropylbenzenesulphonylchloride, EDC = 1-[3-(dimethylamino)propyl]-3-ethylcarbodiimide hydrochloride, MPD = N-methylpyrrolidone, DMF = N,N-dimethylformamide,  $CH_2Cl_2$  = dichloromethane, Py = pyridine, TEAB = triethylammonium bicarbonate, DNP = 2,4-dinitrophenyl,  $NaIO_4$  = sodium m-periodate,  $NaBH_4$  = Sodium borohydride, DIBAL-H = diisobutylaluminiumhydride, THF = tetrahydrofurane,  $K_2CO_3$  = potassium carbonate, EtOH=ethanol, DMT-Cl= dimethoxytrylchloride, R.T.= room temperature, s's = singlets, d's = doublets, m's = multiplets, q's = quartets.

## CHAPTER 5 Synthesis of cIDPR analogues



**Scheme 5.2:** Synthetic scheme followed to obtain compounds III and IV

## 5.11.2. General Procedures

**Synthesis of compounds I and II**

**9-(2',3'-O-isopropylidene -D-ribityl)-hypoxanthine (compound 2).** To a stirred suspension of nucleoside **1** (617 mg, 2 mmol) in anhydrous THF (30 mL) at room temperature was added DIBAL-H (1.0 mL of 1.0M solution in toluene, 10mmol) dropwise under argon atmosphere. After being stirred at 25°C for 24h, the reaction mixture was quenched with saturated 10% aqueous AcOH at 0°C and then evaporated in vacuo. The residue was purified by column chromatography on silica gel (CHCl<sub>3</sub>/MeOH 9:1-8:2), to afford pure **2** in 70% yield (416mg) as a colorless solid.

UV (MeOH)  $\lambda_{\max}$  249nm. <sup>1</sup>H NMR  $\delta_{\text{H}}$  (400MHz: DMSO-d<sub>6</sub>) 1.19 and 1.42 (s, each 3H, isopropylidene), 3.41 (d, 1H, 5'H), 3.61-3.64 (m, 2H, 4'H and 5'H), 4.10 (dd, 1H, 3'H), 4.21 (dd, 1H, 1'H), 4.49 (ddd, 1H, 2'H), 4.53 (dd, 1H, 1'H), 4.64 (t, 1H, 5'OH), 5.12 (d, 1H, 4'OH), 8.02 (s, 2H, H2 and H8), 12.26 (s, 1H, N<sub>1</sub>H). ESI-MS calculated m/z: 310.11, (C<sub>13</sub>H<sub>18</sub>N<sub>4</sub>O<sub>5</sub>), found: 311.50(M + H)<sup>+</sup>.

**4',5'-bis-O-acetyl-9-(2',3'-O-isopropylidene -D-ribityl)-hypoxanthine (3)**

**2** (416 mg, 1.4 mmol) was treated with 10 mL of Ac<sub>2</sub>O/pyridine solution (6:4, v/v). After 2 h at room temperature the dried mixture was dissolved in CHCl<sub>3</sub> and washed with water. The dried organic layer gave 566mg of **3** as white solid (95% yield).

<sup>1</sup>H NMR  $\delta_{\text{H}}$  (400MHz: CHCl<sub>3</sub>-d<sub>1</sub>) 1.19-1.42 (s, each 3H, isopropylidene), 3.67 (s, 6H, acetyl), 3.86 (dd, 1H, 1'H), 3.9 (ddd, 1H, 2'H), 4.40 (d, 2H, 5'H), 4.47 (dd, 1H, 3'H), 4.87 (m, 1H, 4'H), 8.02 (s, 2H, H2 and H8). ESI-MS calculated m/z: 426.14, (C<sub>17</sub>H<sub>22</sub>N<sub>4</sub>O<sub>9</sub>), found: 427.30 (M+H)<sup>+</sup>.

**4',5'-bis-O-acetyl-9-(2',3'-O-isopropylidene -D-ribityl)-N1-nitro-hypoxanthine (4)**

Trifluoroacetic anhydride (350  $\mu$ L, 5.2 mmol) was added to a suspension of finely powdered NH<sub>4</sub>NO<sub>3</sub> (208 mg, 2.6 mmol) in anhydrous CH<sub>2</sub>Cl<sub>2</sub> (5 mL) at 0°C. The mixture was vigorously stirred at room temperature until the solid was dissolved (ca.1h) and then cooled again. Follow the addition of the nucleoside (1.3 mmol, 566 mg) in 3mL CH<sub>2</sub>Cl<sub>2</sub> at -20°C for 60min. The reaction was monitored by TLC and quenched adding cold phosphate buffer. The organic layer dried and then filtered through a pad of silica gel (98:2 CH<sub>2</sub>Cl<sub>2</sub>- MeOH), afforded the pure N1-nitro-derivative **4** as a yellow oil (337mg, 55%yield).

<sup>1</sup>H NMR  $\delta_{\text{H}}$  (400MHz: CHCl<sub>3</sub>-d<sub>1</sub>) 1.19-1.42 (s, each 3H, isopropylidene), 3.67 (s, 6H, acetyl), 3.86 (dd, 1H, 1'H), 3.9 (ddd, 1H, 2'H), 4.40 (d, 2H, 5'H), 4.47 (dd, 1H, 3'H), 4.87 (m, 1H, 4'H), 7.8-8.2 (s, 2H, H2 and H8). ESI-MS calculated m/z: 471.16, (C<sub>17</sub>H<sub>21</sub>N<sub>5</sub>O<sub>11</sub>), found: 472.30(M + H)<sup>+</sup>.

**4',5'-bis-O-acetyl-9-(2',3'-O-isopropylidene-D-ribityl)-N1carbaribosehypoxanthi-ne (6a)**

A solution of **4** (168 mg, 0.35 mmol) in dry DMF (1.0 mL) was treated with **5a** (163 mg, 0.89 mmol) at 80 °C for 2 h under stirring. The resulting solution, dried under reduced pressure, was purified on silica gel column eluted with increasing amount of CH<sub>3</sub>OH in CHCl<sub>3</sub>. The fractions eluted with 12% CH<sub>3</sub>OH furnished 177 mg of **6a** as white amorphous solid (85% yield), which could not be induced to crystallize.

<sup>1</sup>H NMR  $\delta_{\text{H}}$  (400MHz: CHCl<sub>3</sub>-d<sub>1</sub>) 1.42-1.61 (s, each 6H, isopropylidene), 1.69 (dd, 2H, 6''H), 1.91 (dd, 1H, 4''H), 3.49 (d, 2H, 5''H), 3.67 (s, 6H, acetyl), 3.86 (dd, 1H, 1'H), 3.87 (dd, 1H, 3''H), 3.9 (ddd, 1H, 2'H), 3.95 (m, 1H, 1''H), 4.43 (d, 2H, 5'H), 4.40 (dd, 1H, 3'H), 4.50 (dd, 1H, 2''H), 4.87 (m, 1H, 4'H), 7.8-8.2 (s, 2H, H2 and H8). ESI-MS calculated m/z: 596.23, (C<sub>26</sub>H<sub>36</sub>N<sub>4</sub>O<sub>12</sub>), found: 597.30(M + H)<sup>+</sup>.

**4',5'-bis-O-acetyl-9-(2',3'-O-isopropylidene-D-ribityl)-N1-(5-hydroxypenthyI)-hypoxanthine (6b)**

A solution of **4** (168 mg, 0.35 mmol) in dry DMF (1.0 mL) was treated with **5b** (90 mg, 0.89 mmol) at 80 °C for 2 h under stirring. The resulting solution, dried under reduced pressure, was purified on silica gel column eluted with increasing amount



of CH<sub>3</sub>OH in CHCl<sub>3</sub>. The fractions eluted with 15% CH<sub>3</sub>OH furnished 143 mg of **6b** as white amorphous solid (85% yield), which could not be induced to crystallize.

**6b** <sup>1</sup>H NMR δ<sub>H</sub> (400MHz: CH<sub>3</sub>OH-d<sub>3</sub>) 1.41 (s, 3H, isopropylidene), 1.3-1.55 (m, 6H, 2''H, 3''H, 4''H), 2.01 (s, 6H, acetyl) 2.95(t, 1H, 1''H), 3.49 (d, 2H, 5''H), 3.86 (dd, 1H, 1'H), 3.9 (ddd, 1H, 2'H), 4.32 (d, 2H, 5'H), 4.47 (dd, 1H, 3'H), 4.87 (m, 1H, 4'H), 7.97-8.02 (s, 2H, H2 and H8). ESI-MS calculated m/z: 480.22, (C<sub>22</sub>H<sub>32</sub>N<sub>4</sub>O<sub>8</sub>), found: 481.44(M + H)<sup>+</sup>.

**Products 7a and 7b.** Compounds **6a** and **6b** were treated each one with anhydrous K<sub>2</sub>CO<sub>3</sub> in dry MeOH to remove acetyl groups in high yields, affording **7a** and **7b**, respectively.

**7a** <sup>1</sup>H NMR δ<sub>H</sub> (400MHz: CHCl<sub>3</sub>-d<sub>1</sub>) 1.42-1.61 (s, each 6H, isopropylidene), 1.69 (dd, 2H, 6''H), 1.91 (dd, 1H, 4''H), 3.49 (d, 2H, 5''H), 3.62(m, 1H, 4'H), 3.68 (d, 2H, 5'H), 3.86 (dd, 1H, 1'H), 3.87 (dd, 1H, 3''H), 3.9 (ddd, 1H, 2'H), 3.95 (m, 1H, 1''H), 3.96 (dd, 1H, 3'H), 4.50 (dd, 1H, 2''H), 7.8-8.2 (s, 2H, H2 and H8). ESI-MS calculated m/z: 480.23, (C<sub>22</sub>H<sub>32</sub>N<sub>4</sub>O<sub>8</sub>), found: 481.30(M + H)<sup>+</sup>.

**7b** <sup>1</sup>H NMR δ<sub>H</sub> (400MHz: CH<sub>3</sub>OH-d<sub>3</sub>) 1.41(s, 3H, isopropylidene), 1.3-1.55 (m, 6H, 2''H, 3''H, 4''H), 2.95 (t, 1H, 1''H), 3.49 (d, 2H, 5''H), 3.62(m, 1H, 4'H), 3.68 (d, 2H, 5'H), 3.86(dd, 1H, 1'H), 3.9 (ddd, 1H, 2'H), 3.96 (dd, 1H, 3'H), 7.97-8.02 (s, 2H, H2 and H8). ESI-MS calculated m/z: 396.15, (C<sub>18</sub>H<sub>28</sub>N<sub>4</sub>O<sub>6</sub>), found: 397.44(M + H)<sup>+</sup>.

**Products 8a and 8b.** PSS (126 mg, 0.33 mmol) was co-evaporated with dry pyridine (3x2mL) and dissolved in dry Py (2 mL). TPSCI was added (252 mg, 0.66 mmol) and the mixture was stirred at R.T. for 1h. Two solution were prepared, to which one compound **7a** (143 mg, 0.30 mmol, 1eq) and **7b** (118 mg, 0.30 mmol) were separately added. After stirring for 22 h at R.T. the mixtures were diluted with CHCl<sub>3</sub> (10mL) and then water (10mL) was added. The organic layers were dried on Na<sub>2</sub>SO<sub>4</sub>, evaporated to dryness and co-evaporated with toluene. The residues were chromatographed on silica gel with (95:5) CHCl<sub>3</sub>- MeOH, to afford **8a** from **7a** and **8b** from **7b**, respectively in high yields.

**8a** <sup>1</sup>H NMR δ<sub>H</sub> (400MHz: CHCl<sub>3</sub>-d<sub>1</sub>) 1.42-1.61 (s, each 6H, isopropylidene), 1.69 (dd, 2H, 6''H), 1.91 (dd, 1H, 4''H), 3.49 (d, 2H, 5''H), 3.62 (m, 1H, 4'H), 3.68 (d, 2H, 5'H), 3.86 (dd, 1H, 1'H), 3.87 (dd, 1H, 3''H), 3.9 (ddd, 1H, 2'H), 3.95-3.96 (2H, m, 1''H, dd, 3'H), 4.50 (dd, 1H, 2''H), 7.04-7.20 (m, 20H, PhS), 7.8-8.2 (s, 2H, H2 and H8). ESI-MS calculated m/z: 1008.24, (C<sub>46</sub>H<sub>50</sub>N<sub>4</sub>O<sub>10</sub>P<sub>2</sub>S<sub>4</sub>), found: 1009.40(M + H)<sup>+</sup>.

**8b** <sup>1</sup>H NMR δ<sub>H</sub> (400MHz: CHCl<sub>3</sub>-d<sub>1</sub>) 1.41 (s, 3H, isopropylidene), 1.3-1.55 (m, 6H, 2''H, 3''H, 4''H), 2.95 (t, 1H, 1''H), 3.55 (d, 2H, 5''H), 3.62 (m, 1H, 4'H), 3.75 (d, 2H, 5'H), 3.86 (dd, 1H, 1'H), 3.9 (ddd, 1H, 2'H), 3.96 (dd, 1H, 3'H), 7.0-7.2 (m, 20H, PhS), 7.97-8.02 (s, 2H, H2 and H8). ESI-MS calculated m/z: 924.22, (C<sub>42</sub>H<sub>46</sub>N<sub>4</sub>O<sub>8</sub>P<sub>2</sub>S<sub>4</sub>), found: 925.54(M + H)<sup>+</sup>.

**Products 9a and 9b.** One step removal of all four phenylthio-groups, from each **8a** and **8b**, was carried out at R.T. for 16h using 32 equiv. of silver acetate (AgAc) in 10mL of pyridine-water (2:1 v/v). After the deblocking reaction was completed, the excess of AgAc was converted in the less soluble silver chloride, adding NaCl, until solution saturation. Insoluble silver sulfide and silver chloride were separated off by centrifugation. The supernatant was collected, desalted and purified on RP-C18 HPLC to afford **9a** and **9b** from **8a** and **8b**, respectively (90% yield).

**9a** <sup>1</sup>H NMR δ<sub>H</sub> (400MHz: CH<sub>3</sub>OH-d<sub>3</sub>) 1.42-1.61 (s, each 6H, isopropylidene), 1.69 (dd, 2H, 6''H), 1.91 (dd, 1H, 4''H), 3.49 (d, 2H, 5''H), 3.62 (m, 1H, 4'H), 3.68 (d, 2H, 5'H), 3.86 (dd, 1H, 1'H), 3.87 (dd, 1H, 3''H), 3.9 (ddd, 1H, 2'H), 3.95 (m, 1H, 1''H), 3.96 (dd, 1H, 3'H), 4.50 (dd, 1H, 2''H), 7.8-8.2 (s, 2H, H2 and H8). ESI-MS calculated m/z: 640.14, (C<sub>22</sub>H<sub>34</sub>N<sub>4</sub>O<sub>14</sub>P<sub>2</sub>), found: 641.37(M + H)<sup>+</sup>.

**9b** <sup>1</sup>H NMR δ<sub>H</sub> (400MHz: CH<sub>3</sub>OH-d<sub>3</sub>) 1.30-1.55 (m, 6H, 2''H, 3''H, 4''H), 1.42-1.61 (s, 3H, isopropylidene), 1.69 (dd, 2H, 6''H), 2.96 (m, 1H, 1''H), 3.49 (d, 2H, 5''H), 3.62 (m, 1H, 4'H),

3.68 (d, 2H, 5'H), 3.86 (dd, 1H, 1'H), 3.9 (ddd, 1H, 2'H), 3.96 (dd, 1H, 3'H), 7.8-8.2 (s, 2H, H2 and H8). ESI-MS calculated  $m/z$ : 556.24, ( $C_{18}H_{30}N_4O_{12}P_2$ ), found: 557.38( $M + H$ )<sup>+</sup>.

**Products 10a and 10b.** To each solution of **9a** (169 mg, 0.26 mmol) and **9b** (144 mg, 0.26 mmol) in MDP (0.5 mL), EDC (273 mg, 1.43 mmol) was added and the mixture stirred at room temperature for 60 h. Then the mixture was evaporated under reduced pressure and the residue was dissolved in 2 mL of TEAB buffer (0.1 M, pH 7.0). The solution was purified and desalted by a  $C_{18}$  reversed phase HPLC column, to give 133mg of **10a** and 116mg **10b** as bis-triethylammonium salt (80% yield). Spectroscopic data are identical to product **9a** and **9b**.

**Compounds I and II.** A solution of each **10a** and **10b** in aqueous  $HCO_2H$  (60%, 2.0 mL) was stirred at room temperature for 3.5 h. After the solvent was evaporated under reduced pressure and the residue, dissolved in  $H_2O$ , was purified by a  $RP-C_{18}$  HPLC column, eluted with a linear gradient (from 0 to 40% in 80 min) of  $CH_3CN$  in TEAB buffer (0.1 M, pH 7.0), flow 1.0 mL/min. The fractions with retention time of 22.4 min and 23.0 min were evaporated under reduced pressure and excess of TEAB was coevaporated with MeOH. The residue was freeze-dried to give 116 mg of compound **I** and 102mg of **II** as triethylammonium salt, as white amorphous solid (85% yield).

**I from 10a**  $^1H$  NMR  $\delta_H$  (400MHz:  $CH_3OH-d_3$ ) 1.67 (m, 1H, 4''H), 1.69 (dd, 2H, 6''H), 3.30 (dd, 1H, 2''H), 3.33 (ddd, 1H, 2'H), 3.37 (dd, 1H, 3'H), 3.38 (m, 1H, 4'H), 3.49 (d, 2H, 5''H), 3.68 (d, 2H, 5'H), 3.69 (m, 1H, 1''H), 3.88 (dd, 1H, 1'H), 3.94 (dd, 1H, 2''H), 7.8-8.2 (s, 2H, H2 and H8). ESI-MS calculated  $m/z$ : 557.14, ( $C_{16}H_{25}N_4O_{13}P_2$ ), found: 558.37( $M + H$ )<sup>+</sup>.

**II from 10b**  $^1H$  NMR  $\delta_H$  (400MHz:  $CH_3OH-d_3$ ) 1.30-1.55 (m, 6H, 2''H, 3''H, 4''H), 1.69 (dd, 2H, 6''H), 2.96 (m, 1H, 1''H), 3.49 (d, 2H, 5''H), 3.62 (m, 1H, 4'H), 3.68 (d, 2H, 5'H), 3.86 (dd, 1H, 1'H), 3.9 (ddd, 1H, 2'H), 3.96 (dd, 1H, 3'H), 7.8-8.2 (s, 2H, H2 and H8). ESI-MS calculated  $m/z$ : 499.02, ( $C_{15}H_{26}N_4O_{11}P_2$ ), found: 500.15( $M + H$ )<sup>+</sup>.

### Synthesis of compounds III and IV

**5'-(O-dimethoxytrytil)-inosine (compound 2').** To a stirred solution of nucleoside **1'** (308 mg, 1 mmol) in anhydrous Py (5mL) at room temperature DMT-Cl (170 mg, 0.5 mmol) was added. After being stirred at 25°C for 2.5 h, the reaction mixture was quenched with MeOH and then evaporated in vacuo. The residue was purified by column chromatography on silica gel, to afford pure **2'** ( $CHCl_3/MeOH/Pyr$  9:1:0.5) in 90% yield (513 mg) as a yellow solid.  $^1H$  NMR  $\delta_H$  (400MHz: Acetone- $d_6$ ) 3.50 (d, 2H, 5'H), 3.65 (dd, 1H, 3'H), 3.7 (m, 1H, 2'H), 4.15 (m, 1H, 4'H), 6.03 (d, 2H, 1'H), 7.1-7.6 (m, 13H, trytil), 7.97-8.02 (s, 2H, H2 and H8), ESI-MS calculated  $m/z$ : 570.11, ( $C_{31}H_{30}N_4O_7$ ), found: 571.34( $M + H$ )<sup>+</sup>.

#### 5'-(O-dimethoxytrytil)- N1-(2,4-dinitrophenyl)-inosine (compound 3')

A mixture of **2'** (513 mg, 0.9 mmol), 1-chloro-2,4-dinitrobenzene (272 mg, 1.2 mmol) and  $K_2CO_3$  (147 mg, 0.95 mmol) was stirred in anhydrous DMF (5 mL) at 80 °C for 2.5 h. After cooling, the mixture was filtered and the solid was washed with  $CHCl_3$ . The filtrate and washings, evaporated to dryness under reduced pressure, were purified on a silica gel column eluted with increasing amount of  $CH_3OH$  in  $CHCl_3$  (from 0 to 5%) to give 649 mg of **4'** (as 1: 1 mixture of atropoisomers at *N*-1-phenyl bond) as pale yellow amorphous solid (98% yield).  $^1H$  NMR  $\delta_H$  (400MHz: Acetone- $d_6$ ) 3.43 (s, 6H,  $-OCH_3$ ), 3.50 (d, 2H, 5'H), 3.65 (t, 1H, 3'H), 3.7 (m, 1H, 2'H), 4.15 (m, 1H, 4'H), 6.03 (d, 2H, 1'H), 7.1-7.6 (m, 13H, trityl), 7.97-8.12 (s, 2H, H2 and H8), 8.2-9.1 (m, 3H, DNP). ESI-MS calculated  $m/z$ : 736.21, ( $C_{37}H_{32}N_6O_{11}$ ), found: 737.42( $M + H$ )<sup>+</sup>.

**N1-carbaribose-5', 5''-bis-(O-DMT)-inosine (compound 5'a)**

A solution of **4'** (324mg, 0.44 mmol) in dry DMF (2.5 mL) was treated with **a** (196 mg, 1.1 mmol) at 80 °C for 6 h under stirring. The resulting solution, dried under reduced pressure, was purified on silica gel column eluted with increasing amount of CH<sub>3</sub>OH in CHCl<sub>3</sub>. The fractions eluted with 12% CH<sub>3</sub>OH furnished 390 mg of **5'a** as white amorphous solid (85% yield) which could not be induced to crystallize.

<sup>1</sup>H NMR δ<sub>H</sub> (400MHz: Acetone-d<sub>6</sub>) 1.4-1.6 (s, 6H, isopropylidene), 1.7 (dd, 2H, 6''H), 2.15 (m, 1H, 4''H), 3.33 (d, 2H, 5''H), 3.43 (s, 6H, OCH<sub>3</sub>), 4.15 (m, 1H, 4'H), 3.50 (d, 2H, 5'H), 3.87 (t, 1H, 3''H), 3.93 (m, 1H, 1''H), 3.63 (t, 1H, 3'H), 4.50 (m, 1H, 2''H), 3.7 (m, 1H, 2'H), 6.03 (d, 2H, 1'H), 7.1-7.6 (m, 26H, trytil), 7.97-8.12 (s, 2H, H2 and H8). ESI-MS calculated m/z: 1043.47, (C<sub>61</sub>H<sub>62</sub>N<sub>4</sub>O<sub>12</sub>), found: 1044.53(M + H)<sup>+</sup>.

**N1-(5''-ODMT)-penthyl- 5'-(ODMT)-inosine (compound 5'b)**

A solution of **4'** (324 mg, 0.44 mmol) in dry DMF (2.5 mL) was treated with **b** (110 mg, 1.1 mmol) at 80 °C for 6 h under stirring. The resulting solution, dried under reduced pressure, was purified on silica gel column eluted with increasing amount of CH<sub>3</sub>OH in CHCl<sub>3</sub>. The fractions eluted with 15% CH<sub>3</sub>OH furnished 390 mg of **5'b** as white amorphous solid (85% yield) which could not be induced to crystallize.

<sup>1</sup>H NMR δ<sub>H</sub> (400MHz: Acetone-d<sub>6</sub>): 1.3-1.6 (m, 6H, 2''H, 3''H, 4''H), 2.96 (t, 2H, 1''H), 4.15 (m, 1H, 4'H), 3.37 (t, 2H, 5''H), 3.50 (d, 2H, 5'H), 3.65 (dd, 2H, 3'H), 3.7 (m, 2H, 2'H), 6.03 (t, 1H, 1'H), 7.1-7.6 (m, 26H, trytil), 7.97-8.12 (s, 2H, H2 and H8). ESI-MS calculated m/z: 944.20, (C<sub>56</sub>H<sub>56</sub>N<sub>4</sub>O<sub>10</sub>), found: 945.31 (M + H)<sup>+</sup>.

**N1-(5''ODMT)-alkyl-5'-(O-DMT)- 2',3'- seco-inosine (compounds 6'a and 6'b)**

To a stirred solution of each one **5'a** (390 mg, 0.37 mmol) and **5'b** (353 mg, 0.37 mmol) in EtOH/H<sub>2</sub>O (2.5 mL, 1:1 v/v), NaIO<sub>4</sub> ( 119 mg, 0.55 mmol) was added and shaken for 2h at 50°C. At the resulting mixture a solution of NaBH<sub>4</sub> in EtOH ( 0.9mM, 6 mg, 0.5eq) was added dropwise at 0°C and then shaken for 2h at room temperature. The resulting mixture was concentrated under reduced pressure and rapidly filtered on silica gel to afford compound **6'a** and **6'b** (CHCl<sub>3</sub>/MeOH/Pyr 8:2:0.5) in 95% yields as a yellow oil.

**7'a** = <sup>1</sup>H NMR δ<sub>H</sub> (400MHz: Acetone-d<sub>6</sub>) 1.4-1.6(s, 6H, isopropylidene), 1.7(dd, 2H, 6''H), 2.15 (m, 1H, 4''H), 3.25 (m, 1H, 4'H), 3.33 (d, 2H, 5''H), 3.43(s, 6H, OCH<sub>3</sub>), 3.50 (d, 2H, 5'H), 3.87(t, 1H, 3''H), 3.93 (m, 1H, 1''H), 3.66 (d, 1H, 3'H), 4.50 (m, 1H, 2''H), 4.14 (m, 1H, 2'H), 5.13 (d, 2H, 1'H), 7.1-7.6(m, 26H, trytil), 7.97-8.12 (s, 2H, H2 and H8).ESI-MS calculated m/z: 1045.11, (C<sub>61</sub>H<sub>64</sub>N<sub>4</sub>O<sub>12</sub>), found: 1046 (M + H)<sup>+</sup>.

**7'b**= <sup>1</sup>H NMR δ<sub>H</sub> (400MHz: Acetone-d<sub>6</sub>): 1.3-1.6 (m, 6H, 2''H, 3''H, 4''H), 2.96 (t, 2H, 1''H), 4.15 (m, 1H, 4'H), 3.37 (t, 2H, 5''H), 3.50 (d, 2H, 5'H), 3.65 (dd, 2H, 3'H), 3.7 (m, 2H, 2'H), 6.03 (t, 1H, 1'H), 7.1-7.6(m, 26H, trytil), 7.97-8.12(s, 2H, H2 and H8). ESI-MS calculated m/z: 944.20, (C<sub>56</sub>H<sub>58</sub>N<sub>4</sub>O<sub>10</sub>), found: 945.31 (M + H)<sup>+</sup>.

**N1-(5''ODMT)-alkyl-5'-(O-DMT)-2',3'-(bis-O-acetyl)-seco-inosine (compound 6'a and 6'b)**

**5'a** (367 mg, 0.35 mmol) and **5'b** (332 mg, 0.35 mmol) were treated separately with 5 mL of Ac<sub>2</sub>O/pyridine solution (6:4, v/v). After 2 h, at R.T., the stirred mixture was diluted with MeOH, dried under reduced pressure and coevaporated with toluene (10mLx5 times), to give **6'a** and **6'b** as yellow solid (99% yield).

**N1-alkyl-2',3'-(bis-O-acetyl)- seco-inosine (compound 7'a and 7'b)**

Dimethoxytrityl group was removed from **6'a** and **6'b** using a solution 2% TFA in DCM. The mixture was kept for 20 min, evaporated to dryness in vacuo, and coevaporated with pyridine. The residue was chromatographed on silica gel with increasing amount of MeOH in CHCl<sub>3</sub>, affording **7'a** and **7'b** as a white foam.

**7'a**= <sup>1</sup>H NMR δ<sub>H</sub> (400MHz: Acetone-d<sub>6</sub>): 1.4-1.6 (s, 6H, isopropylidene), 1.7 (dd, 2H, 6''H), 2.01 (s, 6H, 2',3'-OCOCH<sub>3</sub>), 2.15 (m, 1H, 4''H), 3.49 (d, 2H, 5''H), 3.52 (m, 1H, 4'H), 3.66 (d, 2H,

5'H), 3.87 (t, 1H, 3''H), 3.93 (m, 1H, 1''H), 4.21 (t, 1H, 3'H), 4.50 (m, 1H, 2''H), 4.67 (m, 1H, 2'H), 5.64 (d, 2H, 1'H), 7.97-8.12 (s, 2H, H2 and H8). ESI-MS calculated m/z: 524.52, ( $C_{23}H_{32}N_4O_{10}$ ), found: 525.76 (M + H)<sup>+</sup>.

**7'b**  $^1H$  NMR  $\delta_H$  (400MHz: Acetone- $d_6$ ): 1.3-1.6 (m, 6H, 2''H, 3''H, 4''H), 2.01 (s, 6H, 2'OCOCH<sub>3</sub>), 2.96 (t, 2H, 1''H), 3.52 (m, 1H, 4'H and 5''H), 3.66 (d, 2H, 5'H), 4.21 (d, 1H, 3'H), 4.69 (d, 1H, 2'H), 5.64 (t, 2H, 1'H), 7.97-8.12 (s, 2H, H2 and H8). ESI-MS calculated m/z: 440.19, ( $C_{19}H_{28}N_4O_8$ ), found: 441.35(M + H)<sup>+</sup>.

**Products 8'a and 8'b.** For procedure see above (analogue products 8a and 8b).

**8'a**  $^1H$  NMR  $\delta_H$  (400MHz: Acetone- $d_6$ ) 1.4-1.6 (s, 6H, isopropylidene), 1.7 (dd, 2H, 6''H), 1.9 (m, 1H, 4''H), 2.01(s, 6H, 2',3'-OCOCH<sub>3</sub>), 3.52-3.6 (m, 3H, 4'H and 5''H), 3.71 (d, 2H, 5'H), 3.87 (t, 1H, 3''H), 3.93 (m, 1H, 1''H), 4.21 (d, 1H, 3'H), 4.50 (m, 1H, 2''H), 4.69 (d, 1H, 2'H), 5.64 (t, 2H, 1'H), 7.04-7.2 (m, 20H, PhS), 7.97-8.12 (s, 2H, H2 and H8). ESI-MS calculated m/z: 1052.17, ( $C_{47}H_{50}N_4O_{12}P_2S_4$ ), found: 1053.15(M + H)<sup>+</sup>.

**8'b**  $^1H$  NMR  $\delta_H$  (400MHz: Acetone- $d_6$ ): 1.3-1.55 (m, 6H, 2''H, 3''H, 4''H), 2.01 (s, 6H, acetyl), 2.96 (t, 2H, 1''H), 3.01 (m, 1H, 4'H), 3.60 (t, 2H, 5''H), 3.73 (d, 2H, 5'H), 4.21 (d, 2H, 3'H), 4.69 (d, 2H, 2'H), 5.64 (t, 1H, 1'H), 7.04-7.2 (m, 20H, PhS), 7.97-8.12 (s, 2H, H2 and H8). ESI-MS calculated m/z: 968.70, ( $C_{43}H_{46}N_4O_{10}P_2S_4$ ), found: 969.90 (M + H)<sup>+</sup>.

**Products 9'a and 9'b.** For procedure see above (analogue products 9a and 9b).

**9'a**  $^1H$  NMR  $\delta_H$  (400MHz: Acetone- $d_6$ ) 1.4-1.6 (s, 6H, isopropylidene), 1.7 (dd, 2H, 6''H), 1.9 (m, 1H, 4''H), 2.01 (s, 6H, 2',3'-OCOCH<sub>3</sub>), 3.52 (m, 3H, 4'H and 5''H), 3.68 (d, 2H, 5'H), 3.73 (m, 1H, 1''H), 3.87 (t, 1H, 3''H), 4.21 (d, 1H, 3'H), 4.50 (m, 1H, 2''H), 4.69 (d, 1H, 2'H), 5.64 (t, 2H, 1'H), 7.97-8.12 (s, 2H, H2 and H8). ESI-MS calculated m/z: 684.14, ( $C_{23}H_{34}N_4O_{16}P_2$ ), found: 685.48(M + H)<sup>+</sup>.

**9'b**  $^1H$  NMR  $\delta_H$  (400MHz: Acetone- $d_6$ ): 1.3-1.55 (m, 6H, 2''H, 3''H, 4''H), 2.01 (s, 6H, acetyl), 2.96 (t, 2H, 1''H), 3.01 (m, 1H, 4'H), 3.58 (t, 2H, 5''H), 3.71 (d, 2H, 5'H), 4.21 (d, 2H, 3'H), 4.69 (d, 2H, 2'H), 5.64 (t, 1H, 1'H), 7.97-8.12 (s, 2H, H2 and H8). ESI-MS calculated m/z: 600.10, ( $C_{19}H_{30}N_4O_{14}P_2$ ), found: 601.16 (M + H)<sup>+</sup>.

**Products 10'a and 10'b.** For procedure see above (analogue products 9a and 9b).

Spectroscopic data are identical to product **9'a** and **9'b**.

**Compound III and IV.** A solution of each **10'a** and **10'b** in anhydrous K<sub>2</sub>CO<sub>3</sub> and dry MeOH was stirred at R.T for 10'. **10'a** requires also the removal step of isopropylidene group, that was performed adding HCO<sub>2</sub>H (60%, 2.0 mL) and stirring at room temperature for 3.5 h. After the solvent was evaporated under reduced pressure and the residue, dissolved in H<sub>2</sub>O, was purified by a RP-C<sub>18</sub> HPLC column, eluted with a linear gradient (from 0 to 40% in 80 min) of CH<sub>3</sub>CN in TEAB buffer (0.1 M, pH 7.0) flow 2.0 mL/min. The fractions with retention time of 20.2 min and 21.5 min were evaporated under reduced pressure and excess of TEAB was coevaporated with H<sub>2</sub>O. Each residue was freeze-dried to give 140 mg of compound **III** and 128 mg of **IV** as triethylammonium salt as white amorphous solid (85% yield).

**III from 10'a**  $^1H$  NMR  $\delta_H$  (400MHz: Acetone- $d_6$ ): 1.7(dd, 2H, 6''H), 1.9 (m, 1H, 4''H), 3.01 (m, 1H, 4'H), 3.49 (d, 2H, 5''H), 3.66 (d, 2H, 5'H), 3.31 (t, 1H, 3''H), 3.66 (d, 1H, 3'H), 3.70 (m, 1H, 1''H), 3.94 (m, 1H, 2''H), 4.14 (d, 1H, 2'H), 5.13 (t, 2H, 1'H), 7.97-8.12 (s, 2H, H2 and H8). ESI-MS calculated m/z: 543.10, ( $C_{16}H_{25}N_4O_{13}P_2$ ), found: 544.36(M + H)<sup>+</sup>.

**IV from 10'b**  $^1H$  NMR  $\delta_H$  (400MHz: Acetone- $d_6$ ): 1.3-1.55 (m, 6H, 2''H, 3''H, 4''H), 2.96 (t, 2H, 1''H), 3.01 (m, 1H, 4'H), 3.53 (t, 2H, 5''H), 3.66 (d, 4H, 3'H and 5'H), 4.14 (d, 2H, 2'H), 5.13 (t, 1H, 1'H), 7.97-8.12 (s, 2H, H2 and H8). ESI-MS calculated m/z: 499.10, ( $C_{15}H_{25}N_4O_{11}P_2$ ), found: 500.16(M + H)<sup>+</sup>.

## References

1. Lee, H. C.; Walset, T. F.; *J.Biol.Chem.*, **1989**, 264, 1608
2. Gu, Q. M.; Sih, C. J.; *J.Am.Chem.Soc.*, **1994**, 116, 7481
3. Dargie, P.G.; Agre, M.C.; *Cell.Regul*, **1990**, 1, 279
4. Galione, A.; Lee, H.C.; *Science*, **1991**, 253, 1143
5. Fleischer, S.; Inui, M.; *Annu.Rev.Biophys.Chem*, **1989**, 18, 333
6. Leichleither, J.; Girard, S.; *Science*, **1991**, 252, 123
7. Currie, K.; Swann, K.; Galione, A.; *Mol. Biol.Cell.*, **1992**, 266, 16985
8. Takasawa, S.; Nata, K.; *Science*, **1993**, 259, 370
9. Munshi, C.; Thiels, D.J. et al.; *J.Biol.Chem.*, **1999**, 274, 30770-30777
10. Lee, H. C.; Prasad, G.S.; *Natur.Struct.Biol.*, **1996**, 3, 957-964
11. Lee, H. C.; Munshi, C.; *J.Biol.Chem.*, **2000**, 275, 21556-21571
12. Sauve, A.A.; Deng, H.T.; *J.Am.Chem.Soc.*, **2000**, 122, 7855-7859
13. Lee, H. C.; Munshi, C.; *J.Biol.Chem.*, **2001**, 276, 12169-12173
14. Deflora, A.; Guida, L.; *FASEB J.*, **1998**, 12, 1507-1520
15. Jacobson, M. K.; Coyle, D. L.; *Methods Enzymol.*, **1997**, 280, 265
16. Kim, H.; Jacobson, E. L.; *Biochem.Biophys.Res.Comm.*, **1993**, 194, 1143.
17. Wada, T.; Inageda, K.; *Nucleosides Nucleotides*, **1995**, 14, 1301.
18. Lee, H. C.; Aarhus, R.; *Nature Struct.Biol.*, **1994**, 1, 143.
19. Wilson, M. H.; McCloskey, J. A. et al.; *J.Org.Chem.*, **1973**, 38, 2247.
20. Ames, B. N.; Martin, R.G.; *J.Biol.Chem.*, **1961**, 236, 2019
21. Kim, H.; Jacobson, E. L.; *Science*, **1993**, 261, 1330
22. Shuto, S.; Fukuoka, M.; *J.Med.Chem*, **2003**, 46, 4741-4749
23. Mort, C. J.; Galione, A.; *Bioorg. Med.Chem*, **2004**, 12, 475-488
24. Muller-Steffner, H., M.; Augustin, H.; *Adv.Exp.Med.Biol*, **1997**, 419, 399.
25. Muller-Steffner, H., M.; Augustin, H.; *Biol.Chem*, **1996**, 271, 23967.
26. Zocchi, E.; Franco, L.; *Biochem.Biophys.Res.Comm.*, **1993**, 196, 1459.
27. Lee, H. C.; Prasad, G. S.; *Proteins*, **1996**, 24, 138.
28. Munshi, C.; Lee, H. C.; *Protein Expr. Purif.*, **1997**, 11, 104
29. Vu, C., Q.; Lu, P., J.; *Biol. Chem*, **1996**, 271, 4747.
30. Zhang, F., J.; Gu, O., M.; *Bioorg. Med. Chem. Lett.*, **1995**, 5, 2267
31. Vu, C., Q.; Coyle, D., L.; *Biochem. Biophys. Res. Commun*, **1997**, 236, 723.
32. Guse, A., H.; Da Silva, C., P.; *Eur. J. Biochem*, **1997**, 245, 411.
33. Schnackerz, K., D.; Vu, C., Q.; *Bioorg. Med. Chem. Lett.*, **1997**, 7, 581.
34. Ashamu, G., A.; Sethi, J., K.; *Biochemistry*, **1997**, 36, 9509.
35. Bailey, V., C.; Fortt, S., M.; *FEBS.Lett.*, **1996**, 379, 227.
36. Walseth, T., F.; Lee, H., C.; *Biochem. Biophys. Acta*, **1993**, 7811, 23
37. Ashamu, G., A.; Galione, A., M.; *J. Chem. Soc. Chem. Commun.*, **1995**, 1359.
38. Zhang, F., J.; Yamada, S.; *Bioorg. Med. Chem. Lett.*, **1996**, 6, 1203.
39. Zhang, F., J.; Sih, C., J.; *Bioorg. Med. Chem. Lett.*, **1996**, 6, 2311.
40. Bailey, V., C.; Fortt, S., M.; *Chem. Biol.*, **1997**, 4, 51.
41. Yamada, S.; Gu, Q., M.; *J. Am. Chem. Soc.*, **1994**, 116, 10787
42. Fortt, S.; Potter B., V., L.; *Tetrahedron Lett.*, **1997**, 38, 5371
43. Shuto, S.; Shirato, M.; *J. Org.Chem.*, **1998**, 63, 1986
44. Shuto, S.; Shirato, M.; Matsuda, A.; *Tetrahedron Lett.*, **1998**, 39, 7341-7344.
45. Galeone, A.; Oliviero, G.; *Tetrahedron*, **2002**, 58, 363-368.
46. Galeone, A.; Oliviero, G.; *Eur. J. Org. Chem.*, **2002**, 4234-4238
47. Wong, L.; Aarhus, R. and Lee, H.C.; *Biochim. Biophys. Acta*, **1999**, 1472, 555
48. Gu, X.; Yang, Z. et al.; *J.Med.Chem*, **2004**, 47, 5674-5682
49. Kumar, A.; Walker, R.T.; *Tetrahedron*, **1990**, 3101-3110
50. Hirota, K.; Monguchi, Y., et al.; *Tetrahedron*, **1997**, 53, 16683-16698
51. Ariza, X.; Vilarrasa, J.; *J.Am.Chem.Soc.*, **1995**, 117, 3665-3673
52. Sekine, M.; Hamaoki, K.; Hata, T.; *Bull.Chem.Soc.Jpn.*, **1981**, 54, 3815-3827

## Appendix

### *List of Publications and Communications:*

1. Amato J., Galeone A., Oliviero G., Mayol L., Piccialli G., Varra M. **Synthesis of 3'-3'-linked pyrimidine oligonucleotides containing an acridine moiety for alternate strand triple helix formation.** European Journal of Organic Chemistry (2004), (11), 2331-2336.
2. Oliviero G., Amato J., Varra M., Piccialli G., Mayol L. **Synthesis of a new N-9 ribityl analogue of cyclic inosine diphosphate ribose (ciDPR) as a mimic of cyclic ADP ribose (cADPR).** Nucleosides, Nucleotides & Nucleic Acids (2005), 24(5-7), 735-738.
3. Oliviero G., Amato J., Borbone N., Galeone A., Varra M., Piccialli G., Mayol, L. **Unusual monomolecular DNA quadruplex structures using bunch-oligonucleotides.** Nucleosides, Nucleotides & Nucleic Acids (2005), 24(5-7), 739-741.
4. Oliviero G., Amato J., Borbone N., Galeone A., Varra M., Piccialli G., Mayol L. **Synthesis and characterization of DNA quadruplexes containing T-tetrads formed by bunch-oligonucleotides.** Biopolymers (2006), 81(3), 194-201.
5. Oliviero G., Amato J., Borbone N., Galeone A., Petraccone L., Varra M., Piccialli G., Mayol L. **Synthesis and Characterization of Monomolecular DNA G-Quadruplexes Formed by Tetra-End-Linked Oligonucleotides.** Bioconjugate Chemistry (2006), 17(4), 889-898.
6. Oliviero G., Borbone N., Amato J., D'Errico S., Piccialli G., Varra M. and Mayol L. **Synthesis of New Ribose Modified Analogues of Cyclic Inosine bis-phosphate Ribose (ciDPR),** in press for Nucleosides, Nucleotides & Nucleic Acids.
7. Borbone N., Mayol L., Miccio L., Oliviero G., Amato J., Pesce G., Piccialli G. and Sasso A.. **Optical Tweezers as a Probe for Oligodeoxyribonucleotides Structuration,** in press for Nucleosides, Nucleotides & Nucleic Acids.
8. Borbone N., Oliviero G., Amato J., D'Errico S., Galeone A., Piccialli G., and Mayol L. **Synthesis and Characterization of Tetra-End Linked Oligonucleotides capable of forming Monomolecular G-Quadruplexes,** in press for Nucleosides, Nucleotides & Nucleic Acids.
9. Oliviero G., Amato J., D'Errico S., Borbone N., Piccialli G. and Mayol L. **Solid phase synthesis of nucleobase and ribose modified inosine nucleoside analogues,** in press for Nucleosides, Nucleotides & Nucleic Acids.
10. Oliviero G., Amato J., Borbone N., D'Errico S., Piccialli G. and Mayol L. **Synthesis of N-1 and ribose modified inosine analogues on solid support,** in press for Tetrahedron Letters.

### *Oral communications*

**Synthesis and Characterization of Monomolecular DNA G-Quadruplexes formed by Tetra-End-Linked Oligonucleotides** XXXI Corso Estivo "A.Corbella" 19/23-06-06

# Synthesis of 3'–3'-Linked Pyrimidine Oligonucleotides Containing an Acridine Moiety for Alternate Strand Triple Helix Formation

Jussara Amato,<sup>[a]</sup> Aldo Galeone,<sup>[a]</sup> Giorgia Oliviero,<sup>[a]</sup> Luciano Mayol,<sup>\*,[b]</sup>  
Gennaro Piccialli,<sup>[a]</sup> and Michela Varra<sup>[b]</sup>

**Keywords:** Acridine / Circular dichroism / Helical structures / Oligonucleotides / Triplex

Oligonucleotides with a 3'–3' inversion of polarity, containing an acridine moiety attached to the nucleotide base flanking the 3'–3' phosphodiester bond, have been synthesised, characterised and used to form alternate-strand

triple helix complexes. These have been investigated by UV melting studies and CD experiments.

(© Wiley-VCH Verlag GmbH & Co. KGaA, 69451 Weinheim, Germany, 2004)

## Introduction

Oligodeoxynucleotides (ODNs) can bind the major groove of double-stranded DNA oligopurine tracts to form local triple helices by sequence-specific hydrogen bonding.<sup>[1]</sup> Pyrimidine-rich, triplex-forming oligonucleotides (TFOs) target, in a parallel orientation, homopurine strands in a duplex DNA by Hoogsteen-type triplets (T•AT and C<sup>+</sup>•GC), while purine-rich TFOs can bind, in an anti-parallel orientation, the purine strand of a duplex DNA by reverse Hoogsteen-type base triplets (G•GC and A•AT).

To extend the range of applicability of this gene-control strategy,<sup>[2]</sup> which is otherwise confined to the rare presence of a long — at least 16–17 bases — homopurine sequence on the target gene, it is possible to form triplex structures in which the duplex is composed of two adjacent and alternating oligopurine-oligopyrimidine tracts. In this approach, called “alternate strand triple helix formation” ODNs having a 3'–3' or 5'–5' internucleoside junction simultaneously hybridise the adjacent purine tracts by switching strand at the junction between the oligopurine and the oligopyrimidine domains of the duplex.<sup>[3]</sup> Several 3'–3' and 5'–5' internucleoside junctions have been proposed as linkers for alternate strand TFOs,<sup>[4–11]</sup> with the latter imparting a minor cooperativity in the binding of the two domains of the TFO with the target duplex.<sup>[5]</sup>

In the last few years we have focused our attention on the synthesis of ODNs containing a 3'–3' inversion po-

larity motif, which is able to hybridize the target duplex, in parallel mode, by formation of Hoogsteen triplets.<sup>[12–15]</sup> Our results on the stability of these triplex structures pointed to the phosphodiester bond as a suitable 3'–3' junction into TFOs.<sup>[12–14,16]</sup>

Nevertheless, molecular-mechanics calculations indicated that a distortion of the sugar-phosphate backbone around the 3'–3' phosphodiester junction occurs, thus preventing the correct Hoogsteen-hydrogen-bond formation for the nucleotide bases flanking the site of inversion polarity.<sup>[17]</sup> In order to balance the destabilizing effects occurring in alternate strand triplexes, several research groups have introduced an intercalating agent near the 3'–3' switching region of the TFO.<sup>[18–20]</sup> Among the possible DNA intercalating agents, acridine is particularly advantageous considering that: i) molecules containing an acridine moiety have been extensively studied as duplex and triplex DNA stabilizing agents; ii) acridine-linked ODNs have been proposed as antisense agents and as efficient TFOs;<sup>[21–23]</sup> iii) ODNs linking acridine derivatives can induce a DNA cleavage under photoirradiation when combined with metal ions;<sup>[24,25]</sup> and iv) a large number of acridine derivatives are commercially available.

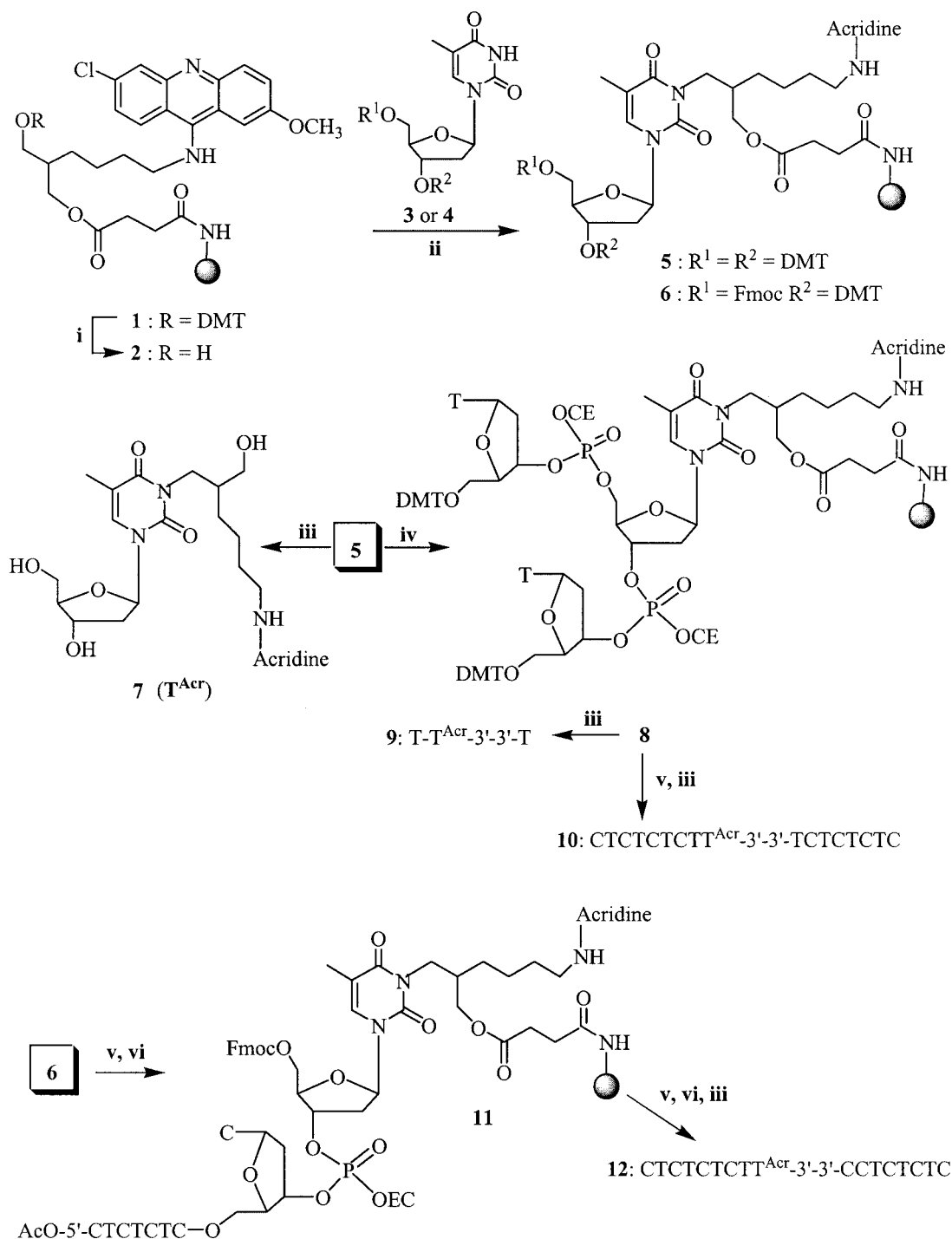
We report here an easy and convenient method for the synthesis of TFOs carrying a chloro-methoxyacridine residue attached to the N-3 atom of a thymidine flanking the 3'–3' phosphodiester bond, through an alkylamino linker. We chose the N-3 position of thymidine on the basis of its inherent reactivity towards alkylation and the reported stability of duplex and triplex DNA complexes containing intercalators, or other groups, attached to the N-3 position of a thymidine residue.<sup>[26,27]</sup>

## Results and Discussion

Our approach is based on a commercially available difunctionalised CPG support carrying a 2-methoxy-6-chlo-

<sup>[a]</sup> Facoltà di Scienze Biotechnologiche, Dipartimento di Chimica delle Sostanze Naturali, Università di Napoli Federico II, Via D. Montesano 49, 80131 Napoli, Italy  
Fax: (internat.) + 39-081-678552  
E-mail: mayoll@unina.it

<sup>[b]</sup> Facoltà di Farmacia, Dipartimento di Chimica delle Sostanze Naturali, Università di Napoli Federico II, Via D. Montesano 49, 80131 Napoli, Italy



Scheme 1. Reagents and conditions: i) TCA 3% in DCM; ii) TPP, DEAD, THF, DCM; iii) NaOH 0.1 M in CH<sub>3</sub>OH/H<sub>2</sub>O (4:1 v/v); iv) coupling with thymidine; v) pyridine/NEt<sub>3</sub> (95:5 v/v, 1.5 h); vi) chain elongation with 3'-phosphoramidite deoxynucleoside building blocks

roacridine moiety and a primary alcoholic hydroxy function protected with the acid-labile DMT group (**1**, Scheme 1).

A suitably 3',5'-protected thymidine **3** or **4** was linked to support **2**, obtained by removal of the DMT protecting group, through the N-3 atom of the base under Mitsunobu condensation conditions,<sup>[28,29]</sup> thus affording supports **5** and **6**, respectively (70% yield by DMT spectrophotometric test). Supports **5** and **6** were found to be stable under the

chemical treatments required by the DNA chain assembly<sup>[30]</sup> and cleavable, at the succinate function, in the final alkaline treatment for detachment and deprotection procedure. The cleavage from the support releases a thymidine base connected to the acridine residue through a hexylamino chain. Support **5**, after capping of the unchanged hydroxy functions and successive removal of the 3',5'-DMT groups, reacted almost quantitatively with the 5'-DMT-thy-



midine-3'-phosphoramidite in a standard automated procedure, giving support **8**. The detritylation of supports **5** and **8**, followed by alkaline treatment with a 0.4 M solution of NaOH, furnished the expected acridine-nucleoside **7** and the acridine-trinucleotide **9**, respectively, the latter containing the 3'-3' phosphodiester junction. The structures of both released products were confirmed by  $^1\text{H}$  NMR spectroscopy and mass spectrometry. Eight coupling cycles performed on support **5**, with appropriate 3'-phosphoramidite nucleotide units, furnished TFO **10**. Support **6**, in which the 5'- and 3'-OH functions are orthogonally protected with Fmoc and DMT groups, respectively, allowed the sequential synthesis of different ODN domains of the TFO. After the removal of the 3'-DMT group from support **6**, the elongation of the first ODN domain by the formation of the 3'-3' phosphodiester junction was achieved, thus leading to **11**. Then, the 5'-Fmoc group was removed by pyridine/ $\text{NEt}_3$  treatment and the second ODN domain was assembled, thus furnishing the TFO **12**. Between the two alternative 5- or 3'-OH Fmoc protections, the former was chosen on the basis of preliminary experiments performed on solid supports carrying 5'- or 3'-Fmoc-thymidine. These have shown that complete 5'-deprotection could be achieved by treatment with 5% triethylamine in pyridine in 1.5 hours, whereas a prolonged reaction time was needed to quantitatively deprotect the 3'-hydroxy function. TFOs **10** and **12** were synthesised on a 1  $\mu\text{M}$  scale by a DNA synthesiser. The crude oligomers, detached from the support as described for trinucleotide **9**, were purified by HPLC on an anionic exchange column. The collected peaks were desalted, thus furnishing **10** and **12** with greater than 95% purity (by RP-HPLC) in 28 and 24  $\text{OD}_{260}$  units yields, respectively. The MALDI TOF-MS spectra of **10** and **12** were in agreement with their structures. Furthermore, the UV/Vis spectra of **10** and **12** confirmed the presence of the acridine moiety in the oligomers, showing a diagnostic absorption at  $\lambda_{\text{max}} = 424 \text{ nm}$ .

### UV and Circular Dichroism (CD) Studies

The triplexes **14** and **15** (Table 1) were formed by mixing equimolar amounts of TFO **10** and **12**, respectively, with

the target duplex DNA **13**, and heating at  $90^\circ\text{C}$  for 5 min. The solutions were then equilibrated for 15 h at room temperature before performing the analyses. The formation and stability of the triplexes **14** and **15**, compared with the acridine-devoid triplexes **16** and **17**, were studied by UV thermal denaturation experiments and by CD in 5 mM  $\text{NaH}_2\text{PO}_4$ , 140 mM KCl and 5 mM  $\text{MgCl}_2$  buffer (at pH 5.5, 6.0 and 6.6). Figure 1 shows the melting profiles of the duplex **13** and triplexes **14**–**17**. The  $T_m$  values are reported in Table 1.

Table 1.  $T_m$  values ( $^\circ\text{C}$ )/pH for duplex **13** and triplexes **14**–**17**

5'-GAGAGAGAATCTCTCTC-3'  
3'-CTCTCTCTTAGAGAGAG-5'

#### Duplex **13**

5'-CTCTCTCTCT-3'  
5'-GAGAGAGAACTCTCTCTC-3'  
3'-CTCTCTCTTAGAGAGAG-5'  
TCCTCTCTC-5'

**Triplex **14****: X = T<sup>Ac</sup>, Y = T

**Triplex **15****: X = T<sup>Ac</sup>, Y = C

**Triplex **16****: X = T; Y = T

**Triplex **17****: X = T; Y = C

Complexes	pH 5.5	pH 6.0	pH 6.6
Duplex <b>13</b>	62.3	63.2	63.4
Triplex <b>14</b>	41.4	36.7	34.5
Triplex <b>15</b>	39.4	34.9	32.5
Triplex <b>16</b>	34.6	29.1	26.0
Triplex <b>17</b>	32.8	27.0	24.6

The melting curves for **14**–**17** show a typical biphasic behaviour, with the first sigmoid attributable to the triplex dissociation, while at higher temperatures the expected pH-insensitive duplex-to-single-strands transition can be observed. The UV melting data for the acridine-containing triplexes **14** and **15** indicate a stabilisation effect at all the pH values with respect to the corresponding acridine-devoid triplexes **16** and **17**. At pH 6.6 the presence of the acridine in the TFO led to a stabilisation, with  $\Delta T_m$  values

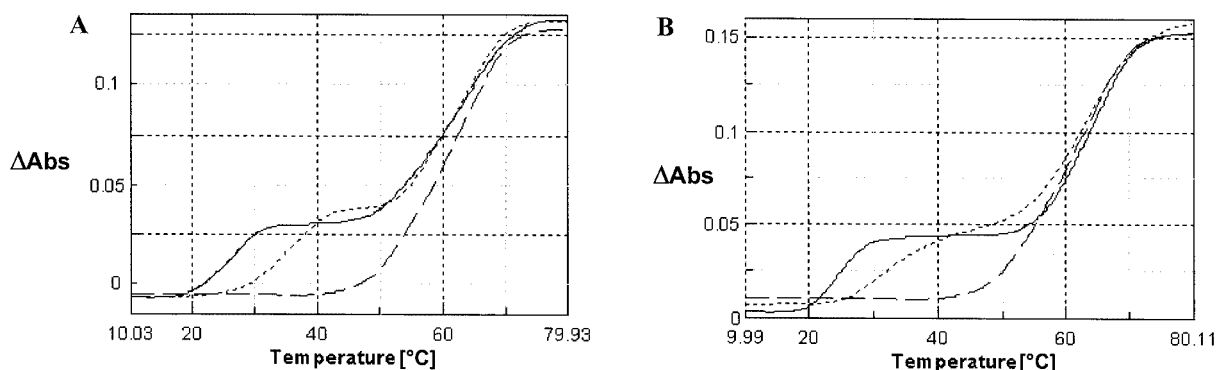


Figure 1. UV melting profiles (pH 6.6): A triplex **14** (····), triplex **16** (—) and duplex **13** (---); B triplex **15** (····), triplex **17** (—) and duplex **13** (---).

of 8.5 and 7.9 °C, respectively, for triplexes **14** and **15**, while at pH 5.5 a slightly smaller effect was observed ( $\Delta T_m$  6.8 and 6.6 °C, respectively). The  $T_m$  values of the triplex **14** were higher than those of **15** at all the tested pH's and the same behaviour was observed for acridine-devoid triplexes **16** and **17**. These latter data indicate that the substitution of a thymine (in **14** and **16**) for a cytosine (in **15** and **17**) at the 3'-3' junction in TFO induces, as expected, a destabilizing effect, probably due to the electrostatic repulsion between consecutive protonated cytosines.

The formation of a triplex structure for **14–17** was confirmed by circular dichroism measurements. The concentration of the triplex solution was in the range  $1.9\text{--}2.3 \times 10^{-5}$  M. The CD spectra of triplexes **14–17** (pH 6.6) are reported in Figure 2. Complexes **16** and **17**, as already reported for the same kind of complexes,<sup>[16,17,31]</sup> show a CD profile very similar to those observed for canonical Hoogsteen-type triplexes, characterised by a large positive band at 280 nm and a negative band centred around 213 nm.<sup>[12,15]</sup> The negative band, in particular, is indicative of the existence of a triplex structure.<sup>[32]</sup> Almost the same patterns were observed for acridine-containing triplexes **14** and **15**. The CD spectra of triplexes **14** and **15** were compared with the normalised summed spectra of the duplex **13** and appropriate single-strand TFOs (Figure 3). Significantly, for

**14** and **15**, the diagnostic band at 213 nm is deeper than the corresponding band in the normalised sum spectra, and the 282 nm band decreases in intensity and shows a red shift. These data are clearly indicative of a triplex structures for complexes **14–17**.

## Conclusion

In this paper we report the easy synthesis of ODNs containing a 3'-3'-phosphodiester linkage and bearing an acridine residue on the thymidine base flanking the 3'-3' junction. This synthesis was based on the preparation of a new kind of nucleoside-acridine solid support (**5** or **6**). In our opinion this synthetic strategy could be a useful entry to link an intercalating agent near to a 3'-3' (or 5'-5') phosphodiester linkage, thus furnishing a convenient method to stabilize alternate strand triple helices by minimizing the mismatching effect into the region between purine and pyrimidine domains. Both CD and UV melting data indicate that the acridine moiety, linked through a seven-atom spacer arm to the N-3 of a thymine, does not hamper the formation of a triplex structure. Furthermore, the stabilisation effect observed for triplexes **14** and **15** strongly suggests an intercalation of the acridine residue into the triplex structure.

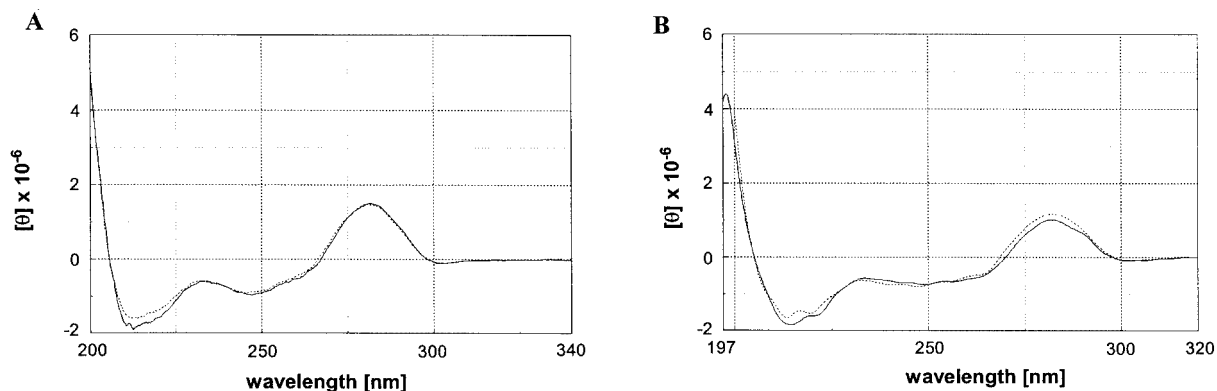


Figure 2. A: CD spectra at pH 6.6 of triplex **14** (---) and triplex **16** (···); B: triplex **15** (---) and triplex **17** (···)

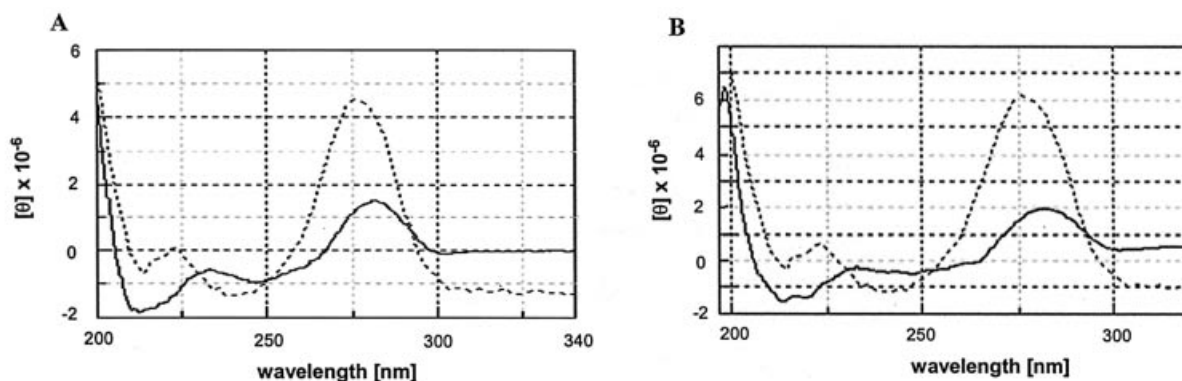


Figure 3. A: CD spectra at pH 6.6 of triplex **14** (---) and normalised summed spectra (···) of duplex **13** and TFO **10**; B: CD spectra at pH 6.6 of triplex **15** (---) and normalised summed spectra (···) of duplex **13** and TFO **12**

## Experimental Section

**General Remarks:** The following abbreviations are used throughout the paper: triplex forming oligonucleotide (TFO), fluorenylmethoxycarbonyl (Fmoc), 4,4'-dimethoxytrityl (DMT), trichloroacetic acid (TCA), dichloromethane (DCM), triethylammonium hydrogen carbonate (TEAB), diethyl azodicarboxylate (DEAD), triphenylphosphane (TPP), tetrahydrofuran (THF).

The acridine CPG support was purchased from Glen Research. Functionalisation of solid supports was carried out in a short column (5 cm length, 1 cm i.d.) equipped with a sintered-glass filter, stopcock and a cap. The ODNs were assembled with a PerSeptive Biosystems Expedite DNA synthesiser using standard phosphoramidite chemistry. HPLC analyses and purification were carried out on a Waters 600E apparatus equipped with a UV detector. UV spectra and thermal denaturation experiments were run with a Jasco V 530 spectrophotometer, equipped with a Jasco 505T temperature controller unit. CD spectra were obtained with a Jasco 715 circular dichroism spectrophotometer. NMR spectra were recorded with a Bruker AMX500 spectrometer. ESI mass spectrometric analyses were performed on a API 2000 (Applied Biosystem) machine used in negative mode. MALDI TOF mass spectrometric analysis was performed on a PerSeptive Biosystems voyager-De Pro MALDI mass spectrometer using picolinic/3-hydroxypicolinic acid mixtures as the matrix.

**3'-O-(4,4'-Dimethoxytriphenylmethyl)-5'-O-fluorenylmethoxycarbonylthymidine (4):**  $^1\text{H}$  NMR ( $[\text{D}_6]$ acetone):  $\delta$  = 9.90 [br. s, 1 H, N(3)-H], 8.02 (s, 1 H, H-6), 7.90–6.80 (m, 21 H's, aromatic protons), 6.32 (m, 1 H, H-1'), 4.52 (m, 2 H,  $\text{CH}_2$  Fmoc residue), 4.32 (m, 1 H, H-3'), 4.24 (t, 1 H, CH Fmoc residue), 3.98 (m, 3 H,  $\text{H}_2$ -5' and H-4'), 3.81 (s, 6 H,  $\text{OCH}_3$  DMT), 1.92 and 1.72 (m, 1 H each,  $\text{H}_2$ -2'), 1.65 (s, 3 H,  $\text{CH}_3$ -5) ppm.

**Preparation of Supports 5 and 6, and Product 7:** Commercially available support **1** (200 mg; 0.044 mequiv./g) was washed with DCM and then treated with  $3 \times 2.0$  mL of a DCM solution of TCA (3%, w/v) for 2 min. This treatment was repeated three times and the resulting support **2** was exhaustively washed with DCM and then dried under reduced pressure.

DEAD (207  $\mu\text{L}$ , 1.3 mmol) was added to a solution of triphenylphosphane (TPP; 340 mg, 1.3 mol) dissolved in 800  $\mu\text{L}$  of THF at 0  $^\circ\text{C}$ . After 10 min this mixture was added to 200 mg of support **2** suspended with a mixture of 0.26 mmol of **5** (or **6**) in 700  $\mu\text{L}$  of anhydrous DCM. After 5 hours at room temperature the support was exhaustively washed with DCM, pyridine and then treated with a solution of pyridine/acetic anhydride (2.0 mL, 9:1 v/v) for 30 min at room temperature. Finally, the support was washed with DCM and  $\text{Et}_2\text{O}$ , and dried under reduced pressure. Incorporation yields of nucleoside **3** and **4** onto **2** were in the range of 65–80% (0.029–0.035 mequiv./g), as determined by a quantitative DMT test performed on dried and weighed samples of the obtained supports **5** and **6**.

Treatment of 50 mg of support **6** with 1.5 mL of a 0.4 M solution of NaOH in methanol/water (4:1, v/v) for 17 h at room temperature, followed by neutralisation of the alkaline solution with a 2 M solution of triethylammonium acetate, furnished product **7** (0.5 mg, after RP HPLC purification).

**7:**  $^1\text{H}$  NMR ( $\text{D}_2\text{O}$ ):  $\delta$  = 8.02, 7.50, 7.48, 7.42, 7.26, 7.18, 7.06 (1 H each, acridine protons and H-6 thymine), 5.56 (m, 1 H, H-1'), 4.60 (m, 1 H, H-3'), 3.90 (s, 3 H,  $\text{CH}_3$ -O), 3.90–3.70 [m's, 5 H, overlapped signals,  $\text{CH}_2$ -NH,  $\text{CH}_2$ -N(1), H-4'], 3.60 (m, 2 H,  $\text{CH}_2\text{OH}$ ), 3.48 (m, 2 H,  $\text{H}_2$ -5'), 2.25 (m, 2 H,  $\text{H}_2$ -2'), 1.61 (s, 3 H,

$\text{CH}_3$ -5), 1.85–1.40 (m's, 7 H, 3  $\text{CH}_2$  and CH) ppm. UV/Vis ( $\text{CH}_3\text{OH}$ ):  $\lambda_{\text{max}}$  = 265 and 422 nm. ESI-MS:  $m/z$  = 613  $[\text{MH}^+]$ .

**Oligonucleotide Synthesis. Products 9, 10 and 12. Symmetric Oligomer Elongation:** Support **5** (50 mg, 0.0016 mequiv.) was used for each synthesis in the automated DNA synthesiser following standard phosphoramidite chemistry, with a 45 mg/mL solution of 3'-phosphoroamidite building block in anhydrous  $\text{CH}_3\text{CN}$ . A commercially available *N*(4)-acetyl-protected cytidine 3'-phosphoramidite building block was used to avoid deamination when sodium hydroxide was used for final cleavage and deprotection. Three-mer **9** and ODN **10** were synthesised by performing one or eight coupling cycles, respectively, with the appropriate nucleotide unit. The coupling yields were always better than 98% (by DMT test). After completion of the desired ODN sequence and final DMT removal, the support was treated with 1.0 mL of a 0.4 M solution of NaOH in methanol/water (4:1, v/v) for 17 h at room temperature. The filtered solution and washings, neutralised with 5.0 mL of a 2 M solution of triethylammonium acetate, were concentrated under reduced pressure, redissolved in water and purified by HPLC.

Crude **9** was purified on an RP  $\text{C}_{18}$  column ( $4.6 \times 250$  mm), eluting with a linear gradient of MeOH in TEAB buffer (pH 7.0, 0 to 10% in 50 min, flow 0.8 mL/min); retention time 35.5 min. The final product (45 units  $\text{OD}_{260}$  from 50 mg of **5**) was lyophilized and characterised.

**9:**  $^1\text{H}$  NMR ( $\text{D}_2\text{O}$ ):  $\delta$  = 7.95, 7.90, 7.85, 7.45, 7.40, 7.32, 7.20, 7.15, 7.02 (9 H, partially overlapped, acridine protons and 3H-6), 6.32–6.40 (m's, 3 H, 3 H-1'), 1.65, 1.68, 1.70 (s's, 3 H each, 3  $\text{CH}_3$ -5) ppm.  $^{31}\text{P}$  NMR ( $\text{D}_2\text{O}$ ):  $\delta$  = 1.55, 1.38 ppm. UV/Vis ( $\text{H}_2\text{O}$ ):  $\lambda_{\text{max}}$  = 265, 422 nm. ESI-MS:  $m/z$  = 1221  $[\text{M} + \text{H}]^+$ , 1243  $[\text{M} + \text{Na}]^+$ .

Crude **10** was purified on a Nucleogel SAX column (Macherey-Nagel, 1000-8/46); buffer A: 20 mM aqueous  $\text{KH}_2\text{PO}_4$  solution, pH 7.0, containing 20% (v/v)  $\text{CH}_3\text{CN}$ ; buffer B: 1.0 M KCl, 20 mM aqueous  $\text{KH}_2\text{PO}_4$  solution, pH 7.0, containing 20% (v/v)  $\text{CH}_3\text{CN}$ ; using a linear gradient from 0 to 100% B in 20 min; flow rate 1.0 mL/min. The collected peak at retention time 18.8 min, desalted on a Sep-Pak column (C18), furnished 28  $\text{OD}_{260}$  units of pure **10**. Maldi TOF-MS (positive mode): calculated mass 5357.03; found 5358.83  $[\text{M} + \text{H}]$ , 5380.86  $[\text{M} + \text{Na}]$ .

**Asymmetric Oligomer Elongation:** Support **6** (50 mg, 0.0015 mequiv.) was used for the synthesis of oligomer **12** following standard phosphoramidite chemistry as described for product **10**. After the elongation of the first ODN tract, the final DMT group was removed and the resulting 5'-OH function capped by reaction with 1 mL of pyridine/ $\text{Ac}_2\text{O}$  solution (4:1, v/v) leading to **11**. Removal of the Fmoc protecting group was achieved by pyridine/ $\text{NEt}_3$  treatment (1.5 h, 95:5, v/v). The resulting 5'-OH support was successively coupled with the required 3'-phosphoramidite derivatives to complete the desired ODN sequence, following a standard automated procedure. Crude **12** was detached from the support, deprotected and purified as described for **10**. The collected peak at retention time 20.5 min, after desalting, furnished 24  $\text{OD}_{260}$  units of pure **12**.

After HPLC purification 24  $\text{OD}_{260}$  units of **12** were obtained. Maldi TOF-MS (positive mode): calculated mass 5342.03; found 5343.65  $[\text{M} + \text{H}]$ , 5383.06  $[\text{M} + \text{K}]$ .

**UV Thermal Denaturation Experiments:** The concentrations of the synthesised oligomers were determined spectrophotometrically at  $\lambda$  = 260 nm and 80  $^\circ\text{C}$ , using the molar-extinction coefficient calculated for unstacked oligonucleotides from the following extinction coefficients: 11700 (G); 8800 (T)  $\text{cm}^{-1} \text{M}^{-1}$ . An aqueous solution of 5 mM  $\text{NaH}_2\text{PO}_4$ , 140 mM KCl and 5 mM  $\text{MgCl}_2$  (at pH 5.5, 6.0

and 6.6) was used for the melting experiments. Melting curves were recorded using a concentration of approximately 1  $\mu\text{M}$  of single strand in 1 mL of the tested solution in Teflon<sup>®</sup>-sealed quartz cuvettes of 1-cm optical path-length. The resulting solutions were then heated at 90 °C for 5 min, then slowly cooled and kept at 20 °C for 15 h. After thermal equilibration at 20 °C the UV absorption at  $\lambda = 260$  nm was monitored as function of the temperature, which was increased at a rate of 0.5 °C/min. The melting temperatures were determined as the maximum of the first derivative of the absorbance vs. temperature plots.

**Circular Dichroism:** CD spectra were registered in the same buffer as used for UV melting experiments at 25 °C in a 0.1 cm path-length cuvette. The wavelength was varied from 200 to 340 nm at 5 nm/min, and the spectra recorded with a response of 16 s, at 2.0 nm bandwidth and normalised by subtraction of the background scan with buffer. The temperature was kept constant at 25 °C with a thermoelectrically controlled cell holder (JASCO PTC-348).

## Acknowledgments

This work was supported by the Italian M.U.R.S.T. (P.R.I.N. 2002 and 2003). The authors are grateful to "Centro Ricerche Interdipartimentale di Analisi Strumentale", C.R.I.A.S., for allowing the use of their NMR facilities.

- [1] N. T. Thuong, C. Hélène, *Angew. Chem. Int. Ed. Engl.* **1993**, *32*, 666–690.
- [2] For a general review, see: B. P. Casey, P. M. Glazer, *Prog. Nucleic Acid Res.* **2001**, *67*, 163–192.
- [3] D. A. Horne, P. B. Dervan, *J. Am. Chem. Soc.* **1990**, *112*, 2435–2437.
- [4] A. Ono, C. N. Chen, L. Kan, *Biochemistry* **1991**, *30*, 9914–9921.
- [5] S. Mc Curdy, C. Moulds, B. C. Froehler, *Nucleosides Nucleotides* **1991**, *10*, 287–290.
- [6] B. C. Froehler, T. Terhorst, J. P. Shaw, S. N. Mc Curdy, *Biochemistry* **1992**, *31*, 1603–1609.
- [7] M. Durand, S. Peloille, N. T. Thuong, J. C. Maurizot, *Biochemistry* **1992**, *31*, 9197–9204.
- [8] P. A. Beal, P. B. Dervan, *J. Am. Chem. Soc.* **1992**, *114*, 4976–4982.
- [9] S. D. Jayasena, B. H. Johnston, *Nucl. Acids Res.* **1992**, *20*, 5279–5288.
- [10] U. Asseline, N. T. Thuong, *Tetrahedron Lett.* **1993**, *34*, 4173–4176.
- [11] U. Asseline, N. T. Thuong, *Tetrahedron Lett.* **1994**, *35*, 5221–5224.
- [12] L. De Napoli, A. Messere, D. Montesarchio, A. Pepe, G. Piccialli, M. Varra, *J. Org. Chem.* **1997**, *62*, 9024–9030.
- [13] L. De Napoli, G. Di Fabio, A. Messere, D. Montesarchio, G. Piccialli, M. Varra, *Eur. J. Org. Chem.* **1998**, 2119–2125.
- [14] L. De Napoli, G. Di Fabio, A. Messere, D. Montesarchio, D. Musumeci, G. Piccialli, *Tetrahedron* **1999**, *55*, 9899–9914.
- [15] G. Barone, L. De Napoli, G. Di Fabio, C. Giancola, A. Messere, D. Montesarchio, L. Petraccone, G. Piccialli, *Bioorg. Med. Chem.* **2001**, *9*, 2895–2900.
- [16] C. Giancola, A. Buono, G. Barone, L. De Napoli, D. Montesarchio, D. Palomba, G. Piccialli, *J. Therm. Anal. Cal.* **1999**, *56*, 1177–1184.
- [17] C. Giancola, A. Buono, D. Montesarchio, G. Barone, *Phys-Chem ChemPhys.* **1999**, *1*, 5045–5049.
- [18] Y. Ueno, M. Mikawa, S. Hoshika, A. Matsuda, *Bioconjugate Chem.* **2001**, *12*, 635–642.
- [19] Y. Ueno, A. Shibata, A. Matsuda, Y. Kitade, *Bioconjugate Chem.* **2003**, *14*, 684–689.
- [20] V. Esposito, A. Galeone, L. Mayol, G. Oliviero, A. Randazzo, M. Varra, *Eur. J. Org. Chem.* **2002**, 4228–4233.
- [21] E. Washbrook, K. R. Fox, *Biochem. J.* **1994**, *301*, 569–575.
- [22] G. C. Silver, C. H. Nguyen, A. S. Boutorine, E. Bisogni, T. Garestier, C. Hélène, *Bioconjugate Chem.* **1997**, *8*, 15–22.
- [23] A. D. Stewart, X. Xu, S. D. Thomas, D. M. Miller, *Nucleic Acid Res.* **2002**, *30*, 2565–2573.
- [24] A. Kuzuya, R. Mizoguchi, F. Morisawa, K. Machida, M. Komiyama, *J. Am. Chem. Soc.* **2002**, *124*, 6887–6894.
- [25] M.-P. Teulade-Fichou, D. Perrin, A. Boutorine, D. Polverari, J.-P. Vigneron, J.-M. Lehn, J.-S. Sun, T. Garestier, C. Hélène, *J. Am. Chem. Soc.* **2001**, *123*, 9283–9292.
- [26] M. Kubota, K. Ono, *Nucl. Acids Res. Suppl.* **2002**, *2*, 33–34.
- [27] K. Ono, Jpn. Kokai Tokkyo Koho 2003, 8 pp. CODEN: JKXXXAF JP 2003088374 A2 20030325 CAN 138:255457 AN 2003:227376 CAPLUS.
- [28] O. Mitsunobu, *Synthesis* **1981**, 1–28.
- [29] M. de Champdoré, L. De Napoli, G. Di Fabio, A. Messere, D. Montesarchio, G. Piccialli, *Chem. Commun.* **2001**, 2598–2599.
- [30] *Oligonucleotides and Analogues: A Practical Approach* (Ed.: F. Eckstein), IRL Press, Oxford, UK, **1991**.
- [31] L. M. Xodo, G. Manzini, F. Quadrifoglio, *Nucleic Acids Res.* **1990**, *18*, 3557–3564.
- [32] V. P. Antao, D. M. Grey, R. L. Ratliff, *Nucleic Acids Res.* **1988**, *16*, 719–783.

Received December 22, 2003



## SYNTHESIS OF A NEW N-9 RIBITYL ANALOGUE OF CYCLIC INOSINE DIPHOSPHATE RIBOSE (cIDPR) AS A MIMIC OF CYCLIC ADP RIBOSE (cADPR)

Giorgia Oliviero, Jussara Amato, Michela Varra, Gennaro Piccialli, and Luciano Mayol □ *Dipartimento di Chimica delle Sostanze Naturali, Università di Napoli Federico II, Napoli, Italy*

□ *A new analogue of cyclic inosine diphosphate ribose (cIDPR), in which the N-1 and N-9 ribosyl moieties were substituted by a carbocyclic moiety and a hydroxyl-alkyl chain, has been synthesized and characterized.*

**Keywords** Cyclic ADP Ribose (cADPR), cIDPR, IP<sub>3</sub>

### INTRODUCTION

Many cellular functions are modulated by the concentration of intracellular calcium ions.<sup>[1]</sup> Two major mechanisms of calcium mobilization are known that utilize calcium stored in cytoplasmic compartments for signalling. The release of Ca<sup>2+</sup> is triggered by the interaction of the second messenger, inositol 1,4,5-triphosphate (IP<sub>3</sub>), with its receptor, a ligand-activated calcium-selective channel. A second class of intracellular calcium-releasing channels is the ryanodine receptor.<sup>[2]</sup> Although the physiological activator of this receptor is unknown, it can be activated by Ca<sup>2+</sup>, causing the so called Ca<sup>2+</sup>-induced Ca<sup>2+</sup> release (CICR). Recent research establishes that, in addition to inositol triphosphate, the internal calcium stores can be mobilized by new messenger molecules *via* cyclic ADP-ribose (cADPR, Figure 1). cADPR that serves as a second messenger to activate the ryanodine receptors of the sarcoplasmic reticulum and to mobilize intracellular Ca<sup>2+</sup> in many cell types in different species covering protozoa, plants, animals, and human.<sup>[3]</sup> However, the mechanisms mediating the effect of cADPR remain unknown. Since Ca<sup>2+</sup> ions are regulators of several cell functions, muscle contraction, secretion of neurotransmitters, hormones and enzymes, fertilization of oocytes, and

Address correspondence to Giorgia Oliviero, Dipartimento di Chimica delle Sostanze Naturali, Università di Napoli Federico II, Napoli, Italy; E-mail: [golivier@unina.it](mailto:golivier@unina.it)

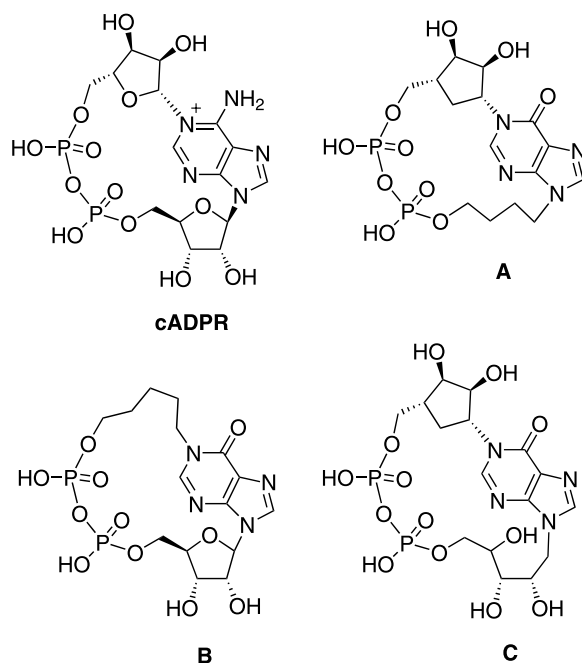


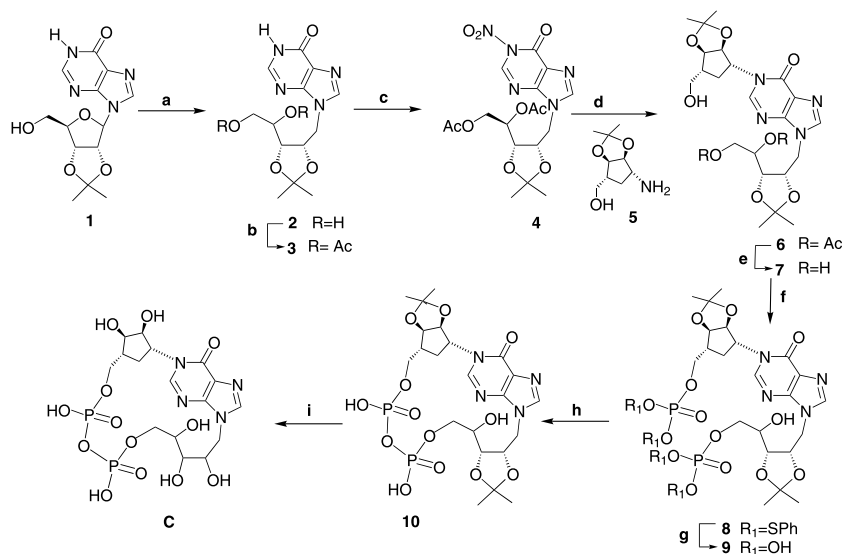
FIGURE 1

lymphocyte activation and proliferation, the cADPR signaling pathway may become a valuable target for pharmaceutical intervention.

cADPR is characterized by a very labile *N*-1 glycosidic bond which is rapidly hydrolyzed both enzymatically, by cADP hydrolase, and non-enzymatically, to give ADP-ribose even in neutral aqueous solution. This biological and chemical instability deeply hinders further studies on cADPR aimed at elucidating its physiological role, particularly as far as regulation of  $\text{Ca}^{2+}$  mobilization in the cells is concerned. Hence, stable yet active cADPR analogues are of great interest in this field.<sup>[4]</sup>

In our laboratories we have previously synthesized two analogues of cADPR (**A** and **B**, Figure 1).<sup>[5,6]</sup> These compounds, which display a carboribosyl and a butyl moiety at N-1 and N-9 of a hypoxanthine base, respectively, are designed to investigate the role of a ribosyl moiety in the mechanism of  $\text{Ca}^{2+}$  intracellular modulation. We wish to report here the synthesis of a new cADPR analogue (**C**, Figure 1) having a three-hydroxylated butyl chain at the N-9 of the hypoxanthine base and a carboribosyl moiety at the N-1 position, which is expected to be resistant to both enzymatic and chemical hydrolysis as are the similar N-1 alkylated analogues.

The synthetic strategy adopted for **C** is shown in Scheme 1. The 2',3'-O-isopropylidene-inosine **1** was allowed to react with DIBAL-H in anhydrous THF. In this way the sugar moiety was reduced to afford the corresponding 1-D-ribitylinosine derivative **2**. Acetylation of the hydroxyl functions of **2** gave derivative **3**. N-1



**SCHEME 1** Reagents and conditions: a) DIBAL-H, THF; b) Ac<sub>2</sub>O, pyridine, r.t.; c) NH<sub>4</sub>NO<sub>3</sub>, TFA, CH<sub>2</sub>Cl<sub>2</sub>, 0°C; d) **5**, DMF, r.t.; 8 h; e) K<sub>2</sub>CO<sub>3</sub>, CH<sub>3</sub>OH, r.t.; 15 min; f) S,S-diphenylphosphorodithioate, TPSCl, pyridine, N<sub>2</sub>, r.t.; 8 h; g) AcAg, pyridine/H<sub>2</sub>O; h) EDC, NMP, r.t.; 60 h; i) aq. CH<sub>3</sub>CO<sub>2</sub>H, r.t.; 3.5 h.

nitroinosine derivative **4** was prepared according to the procedure proposed by Vilarrasa and coworkers<sup>[7]</sup> by treatment of **3** with ammonium nitrate and tri-fluoroacetic anhydride (TFAA) in dichloromethane. Derivative **6** could be obtained by reaction of **4** with **5** in DMF as previously described.<sup>[5]</sup> Removal of the acetyl protecting groups of **6** by treatment with K<sub>2</sub>CO<sub>3</sub> in MeOH yielded **7**. The bis(phenylthio)phosphoryl groups were introduced on the primary hydroxyl functions by treatment of **7** with S,S-diphenyl-phosphorodithioate (cyclohexylammonium salt) in the presence of 2,4,6-triisopropylbenzenesulphonylchloride (TPSCl) and tetrazole in pyridine to give the protected bisphosphate derivative **8** as the main product together with the triphosphate derivative in a 8:2 ratio. After HPLC (RP18) purification, **8** was treated with silver acetate (Ag Ac) for complete removal of protecting groups of both phosphate residues. The deprotected bisphosphate derivative **9** was further purified by HPLC (RP18) before cyclization, which was performed as previously described,<sup>[6]</sup> to obtain **10** (50% yield after purification). Finally, **10** was deprotected at phosphate residues by treatment with acetic acid, which afforded the target compound **C** (10% overall yield starting from **1**) whose structure was confirmed by <sup>1</sup>H and <sup>31</sup>P NMR and MS data.

## REFERENCES

- Guse, A.H. Biochemistry, biology, and pharmacology of cyclic adenosine diphosphoribose (cADPR). *Cur. Med. Chem.* **2004**, *11*, 847–855.
- Zhang, F.J.; Gu, Q.M.; Sih, C.J. Bioorganic chemistry of cyclic ADP-ribose (cADPR). *Bioorg. Med. Chem.* **1999**, *7*, 653–662.

3. Barata, H.; Thompson, M.; Zielinska, W.; Han, Y.S.; Mantilla, C.B.; Prakash, Y.S.; Feitoza, S.G.; Chini, E.N. The role of cyclic-ADP-ribose-signaling pathway in oxytocin-induced  $\text{Ca}^{2+}$  transients in human myometrium cells. *Endocrinology* **2004**, *145*, 881–892.
4. Shuto, S.; Matsuda, A. Chemistry of cyclic ADP-ribose and its analogs. *Cur. Med. Chem.* **2004**, *11*, 827–845.
5. Galeone, A.; Mayol, L.; Oliviero, G.; Piccialli, G.; Varra, M. Synthesis of a novel N-1 carbocyclic, N-9 butyl analogue of cyclic ADP ribose (cADPR). *Tetrahedron* **2002**, *58*, 363–368.
6. Galeone, A.; Mayol, L.; Oliviero, G.; Piccialli, G.; Varra, M. Synthesis of a new N-1-pentyl analogue of cyclic inosine diphosphate ribose (cIDPR) as a stable potential mimic of cyclic ADP ribose (cADPR). *Eur. J. Org. Chem.* **2002**, *24*, 4234–4238.
7. Ariza, X.; Bou, V.; Vilarrasa, J. A new route to  $^{15}\text{N}$ -labeled, N-alkyl, and N-amino nucleosides via N-nitration of uridines and inosines. *J. Am. Chem. Soc.* **1995**, *117*, 3665–3673.



Copyright of Nucleosides, Nucleotides & Nucleic Acids is the property of Taylor & Francis Ltd. The copyright in an individual article may be maintained by the author in certain cases. Content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.

## UNUSUAL MONOMOLECULAR DNA QUADRUPLIX STRUCTURES USING BUNCH-OLIGONUCLEOTIDES

**Giorgia Oliviero, Jussara Amato, Nicola Borbone, Aldo Galeone, Michela Varra, Gennaro Piccialli, and Luciano Mayol** □ *Dipartimento di Chimica delle Sostanze Naturali, Università di Napoli Federico II, Napoli, Italy*

□ *The chemical synthesis of several G-rich bunch-oligonucleotides and the structural characterization of the corresponding monomolecular G-quadruplexes (I–IV) have been reported. The synthetic method allow the achievement of monomolecular DNA quadruplex structures having unusual and predeterminable oligodeoxyribonucleotide (ODN) strand orientation.*

**Keywords** G-Quadruplex, Bunch-ODN, Solid-Phase Synthesis

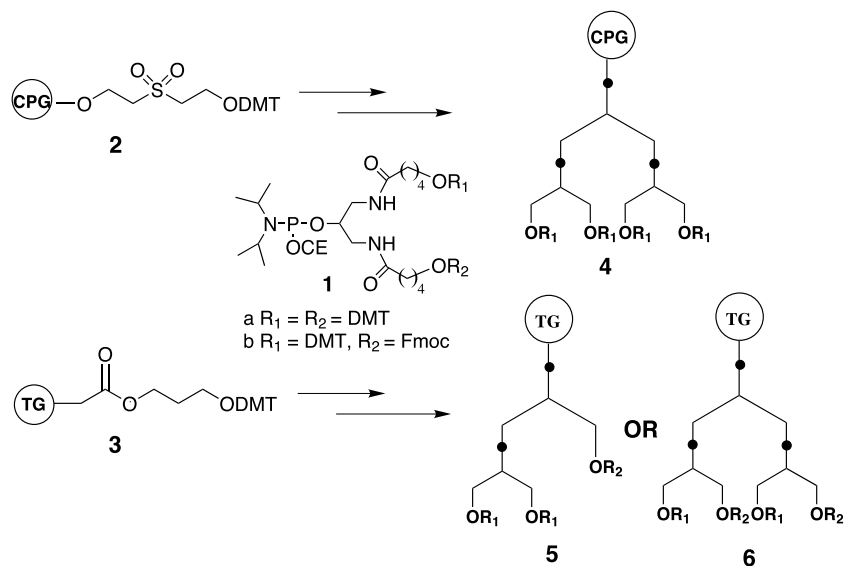
### INTRODUCTION

DNA quadruple helices based on G-quartets (G-quadruplexes) have aroused widespread interest not only for their substantiated presence in many biologically important regions of the genome but also because such structures form the scaffold of several aptamers provided with useful biological properties.<sup>[1–6]</sup> G-quadruplexes can be classified on the basis of the number of self-associating strands one, two, or four strand and are further distinguished by the orientation of the strands to each others (parallel or antiparallel).<sup>[7]</sup>

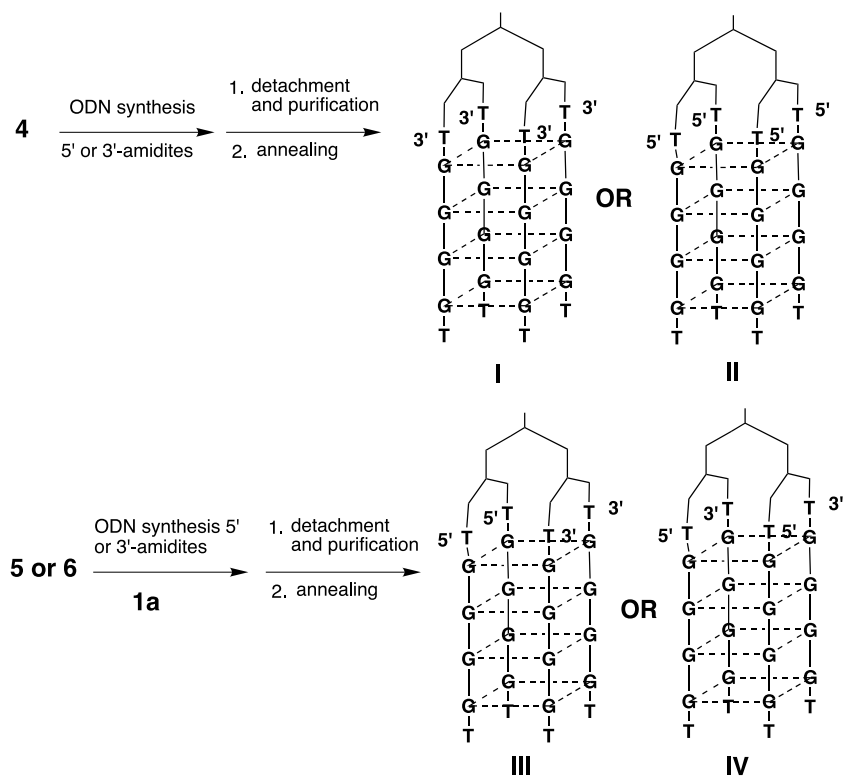
The monomolecular complexes, largely involved in biomolecular events are, almost exclusively, of antiparallel type, showing a higher stability than bi- or tetramolecular counterparts. On the contrary, tetramolecular complexes, whose formation is characterized by unfavorable kinetic and thermodynamic parameters, show the four strands in a parallel orientation. Generally, only the quadruplexes having an adequate stability are suitable for structural investigations.

In this frame the achievement of stable quadruplex models, having predeterminable strand orientation or less stable quartets could be useful for biological and structural studies. The intermolecular formation of parallel structures in vitro is very slow and may require high ODN concentrations. These unfavorable kinetic

Address correspondence to Giorgia Oliviero, Dipartimento di Chimica delle Sostanze Naturali, Università di Napoli Federico II, Napoli, Italy; E-mail: [golivier@unina.it](mailto:golivier@unina.it)



**SCHEME 1** General synthetic procedure for the functionalization of the solid supports **2** and **3** with the bunch-spacers **1a** or **1b**.



**SCHEME 2** Synthetic procedure for the quadruplexes **I**–**IV**.

and thermodynamic parameters could be disadvantageous in view of their potential therapeutic use.

We recently synthesized a new class of ODNs analogues, which we called bunch-ODNs,<sup>[8]</sup> capable to form very stable monomolecular G-quadruplex structures. The structural feature of these analogues is the presence of four ODN strands whose 3'-ends are linked together by a tetra-branched spacer (bunch-spacer). Our solid-phase synthetic strategies uses a commercially available bifunctional linker (**1a–b**, Scheme 1) having symmetrical or orthogonal protected alcoholic functions. As solid support we employed the commercially available CPG-resin **2** or a suitable functionalized carboxy-TentaGel-resin **3** which allow the release of the bunch-ODNs by mild basic treatment. Using a tailored synthetic pathway (Scheme 2), which uses nucleotides 3' or 5' phosphoramidite building blocks, the bunch-quadruplexes d[(TGGGGT)]<sub>4</sub> **I–IV** having a predetermined strands orientation, were obtained.

The correct structure of the bunch-ODNs was ascertained by <sup>1</sup>H NMR and MALDI data. Furthermore, the structure and the stability of the G-quadruplexes (**I–IV**) were investigated by CD thermal denaturation and <sup>1</sup>H NMR experiments at variable temperature. Preliminary results indicate that all bunch-ODNs are capable to adopt a G-quadruplex structure and that the bunch-spacer leads to a more stable complex (**I**) when linked to 3'-ODN ends.

## REFERENCES

1. Shafer, R.H.; Smirnov, I. Biological aspects of DNA/RNA quadruplexes. *Biopolymers* **2001**, *56*, 209–227.
2. Arthanari, H.; Bolton, P.H. Functional and dysfunctional roles of quadruplex DNA in cells. *Chem. Biol.* **2001**, *8*, 221–230.
3. Jing, N.; Marchand, C.; Guan, Y.; Liu, J.; Pallansch, L.; Lackman-Smith, C.; De Clercq, E.; Pommier, Y. Structure-activity of inhibition of HIV-1 integrase and virus replication by G-quartet oligonucleotides. *DNA Cell Biol.* **2001**, *20*, 499–508.
4. Smirnov, I.; Shafer, R.H. Effect of loop sequence and size on DNA aptamer stability. *Biochemistry* **2000**, *39*, 1462–1468.
5. Pileur, F.; Andreola, M.-L.; Dausse, E.; Michel, J.; Moreau, S.; Yamada, H.; Gaidamarov, S.A.; Crouch, R.J.; Toulmé, J.J.; Cazenave, C. Selective inhibitory DNA aptamers of the human RNase H1. *Nucleic Acids Res.* **2003**, *31*, 5776–5788.
6. Chinnappen, D.J.F.; Sen, D. Hemin-stimulated docking of cytochrome c to a hemin-DNA aptamer complex. *Biochemistry* **2002**, *41*, 5202–5212.
7. Davis, J.T. G-quartets 40 years later: from 5'-GMP to molecular biology and supramolecular chemistry. *Angew. Chem., Int. Ed.* **2004**, *43*, 668–698.
8. Oliviero, G.; Borbone, N.; Galeone, A.; Varra, M.; Piccialli, G.; Mayol, L. Synthesis and characterization of a bunchy oligonucleotide forming a monomolecular parallel quadruplex structure in solution. *Tetrahedron Lett.* **2004**, *45*, 4869–4872.

Copyright of Nucleosides, Nucleotides & Nucleic Acids is the property of Taylor & Francis Ltd. The copyright in an individual article may be maintained by the author in certain cases. Content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.

Giorgia Oliviero

Jussara Amato

Nicola Borbone

Aldo Galeone

Michela Varra

Gennaro Piccialli

Luciano Mayol

Dipartimento di Chimica delle  
Sostanze Naturali, Università  
degli Studi di Napoli "Federico  
II," Via D. Montesano 49,  
I-80131 Napoli, Italy

Received 30 July 2005;

revised 20 September 2005;

accepted 10 October 2005

Published online 18 October 2005 in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/bip.20399

# Synthesis and Characterization of DNA Quadruplexes Containing T-Tetrads Formed by Bunch-Oligonucleotides

**Abstract:** The solid phase syntheses of the bunch oligonucleotides **1** and **2** based on the sequences of the natural oligodeoxynucleotides (ODNs)  $d(TG_2TG_2C)$  (**3**) and  $d(CG_2TG_2T)$  (**4**), respectively, attached to a non-nucleotidic tetrabranched linker, are reported. Bunch-ODNs **1** and **2** were shown to form more stable monomolecular parallel G-quadruplexes **I** and **II** when compared with their tetramolecular counterparts  $[d(TG_2TG_2C)]_4$  (**III**) and  $[d(CG_2TG_2T)]_4$  (**IV**), respectively. The structure and stability of all the synthesized complexes have been investigated by circular dichroism (CD), CD thermal denaturation experiments, and  $^1H$ -NMR (nuclear magnetic resonance) experiments at variable temperatures. Particularly, the spectroscopic data confirmed that **1** adopts a T-tetrad containing parallel-stranded quadruplex structure **I** as in the tetramolecular complex **III**.

© 2005 Wiley Periodicals, Inc. *Biopolymers* 81: 194–201, 2006

This article was originally published online as an accepted preprint. The "Published Online" date corresponds to the preprint version. You can request a copy of the preprint by emailing the *Biopolymers* editorial office at [biopolymers@wiley.com](mailto:biopolymers@wiley.com)

**Keywords:** DNA; T-tetrad; oligonucleotide; quadruplex; bunch-ODN

## INTRODUCTION

Quadruplex structures are the main topic of a very large number of reports dealing with their chemistry, molecular biology, and recently, pharmacologic implications.<sup>1</sup> These structures are likely to form in G-rich sequences and are characterized by a stacked arrangement of several GGGG planar tetrads in which each G interacts with the adjacent one via two hydrogen

bonds and behaves both as acceptor and donor of what is known as a Hoogsteen base pair.

Although G-rich sequences occur very frequently in the human genome<sup>2,3</sup> and the formation of G-quadruplexes in vitro is well established, direct evidence for their existence in vivo is just beginning to appear.<sup>4–6</sup> In contrast, many reports have shown the presence of G-quadruplexes in several biologically active oligodeoxynucleotides (ODNs) known as aptamers.<sup>7,8</sup>

Correspondence to: Gennaro Piccialli; e-mail: [piccialli@unina.it](mailto:piccialli@unina.it)  
*Biopolymers*, Vol. 81, 194–201 (2006)  
© 2005 Wiley Periodicals, Inc.

In spite of their common features in the sequences (the need of several G-tracts), G-quadruplex structures so far described showed an unexpected structural variability. In fact, quadruplex structures can differ in the stoichiometry of strands, their relative orientation, the glycosidic torsion angle of each G residue, the nature of the coordinating cation, and both the sequence and size of spacer stretches separating guanine tracts. The occurrence of participation of both natural and modified nonguanine nucleotides in tetrad formation further increased the variety of quadruplex structures. As a matter of fact, several unusual tetrad such as A-tetrads,<sup>9–11</sup> T-tetrads,<sup>12,13</sup> C-tetrads,<sup>14</sup> mixed tetrads (for a recent example, see Ref. 15), and modified bases containing tetrads<sup>16,17</sup> have been described.

Among these, T-tetrads (Figure 1) have aroused our interest since methyl groups of T residues could give rise to hydrophobic contacts into G-quadruplex grooves potentially useful for improving the interaction between quadruplex-based aptamers and their target molecules.

Even though the presence of a T-tetrad in quadruplex structures could be an attractive topic, particularly in the aptamer research area, it shows a main drawback due to the minor stability of a T-tetrad (only four hydrogen bonds and a smaller surface for stacking interactions in comparison with a G-tetrad) that could decrease the stability of the overall complex. Recently, our group has carried out the synthesis and structural characterization of a bunch-ODN able to form a monomolecular parallel quadruplex structure—namely, bunch-[d(TG<sub>4</sub>T)]<sub>4</sub>, in which 3'-ends are linked together by a tetrabranching spacer.<sup>18</sup> Since this ODN analogue is provided with more favorable thermodynamic properties than its tetramolecular counterpart, we reasoned that the destabilization of a quadruplex structure due to the presence of a T-tetrad could be counterbalanced by the improvement caused by the bunchy spacer.

In order to verify this hypothesis, we report herein the synthesis of the [d(TG<sub>2</sub>TG<sub>2</sub>C)]<sub>4</sub>-bunch (**1**) (Scheme 1) whose base sequence matches that of the

ODN **3** forming the tetramolecular quadruplex structure **III** in which a T-tetrad was observed for the first time.<sup>12</sup> The resulting bunch-quadruplex **I** was investigated by circular dichroism (CD), CD thermal denaturation experiments, <sup>1</sup>H, and nuclear Overhauser effect spectroscopy (NOESY) nuclear magnetic resonance (NMR) experiments. Further, we report our studies on the bunch-[d(CG<sub>2</sub>TG<sub>2</sub>T)]<sub>4</sub> (**2**) and its corresponding natural ODN **4**, in which, in respect to **1**, the C and T end bases are mutually exchanged, with the aim to explore the capability of the tetrabranching linker to stabilize other sequences potentially prone to form quadruplex structures containing a T-tetrad.

## RESULTS AND DISCUSSION

### Synthesis and Purification of Quadruplexes I–IV

The syntheses of bunch-ODNs **1** and **2** (Scheme 1) have been performed using the recently reported solid-phase procedure,<sup>18</sup> which uses the hydroxyl-functionalized Controlled Pore Glass (CPG) support **5** and the bifunctional linker **6** to prepare support **7** bearing a symmetrical tetrabranching linker with four protected primary alcoholic functions, prone to ODN-chain assembly. Support **7** was used to synthesize the bunch-ODNs **1** and **2**, each characterized by having the four ODN strands linked to the spacer through the 3'- or 5'-end, respectively. The ODN chains were assembled by an automatic DNA synthesizer using 3'- or 5'-phosphoramidite nucleotide building blocks, thus obtaining the polymer bound bunch-ODNs **8** and **9** respectively. Supports **8** and **9** treated with conc. NH<sub>4</sub>OH (7 h, 55°C), afforded crude bunch-ODNs **1** and **2**, which after high performance liquid chromatography (HPLC) purification and annealing procedure furnished the quadruplexes **I** and **II**, respectively. The natural ODNs **3** and **4**, precursors of the tetramolecular quadruplexes **III** and **IV**, respectively, have been prepared using a standard automatic DNA synthetic procedure followed by HPLC purification. The purity of the products was checked by polyacrylamide gel electrophoresis (PAGE) and the structures confirmed by <sup>1</sup>H-NMR and matrix-assisted laser desorption ionization–mass spectroscopy (MALDI-MS) data. The quadruple helices **I–IV** were obtained dissolving the samples **1–4** in K<sup>+</sup> buffers and annealing to assure the correct formation of quadruplex structures.

### CD Experiments

To investigate whether bunch-ODNs **1** and **2** adopt the quadruplex structures **I** and **II**, and to determine

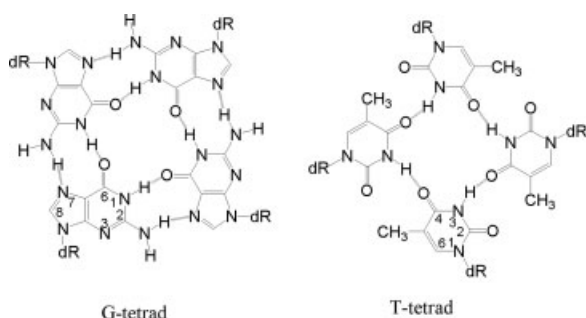
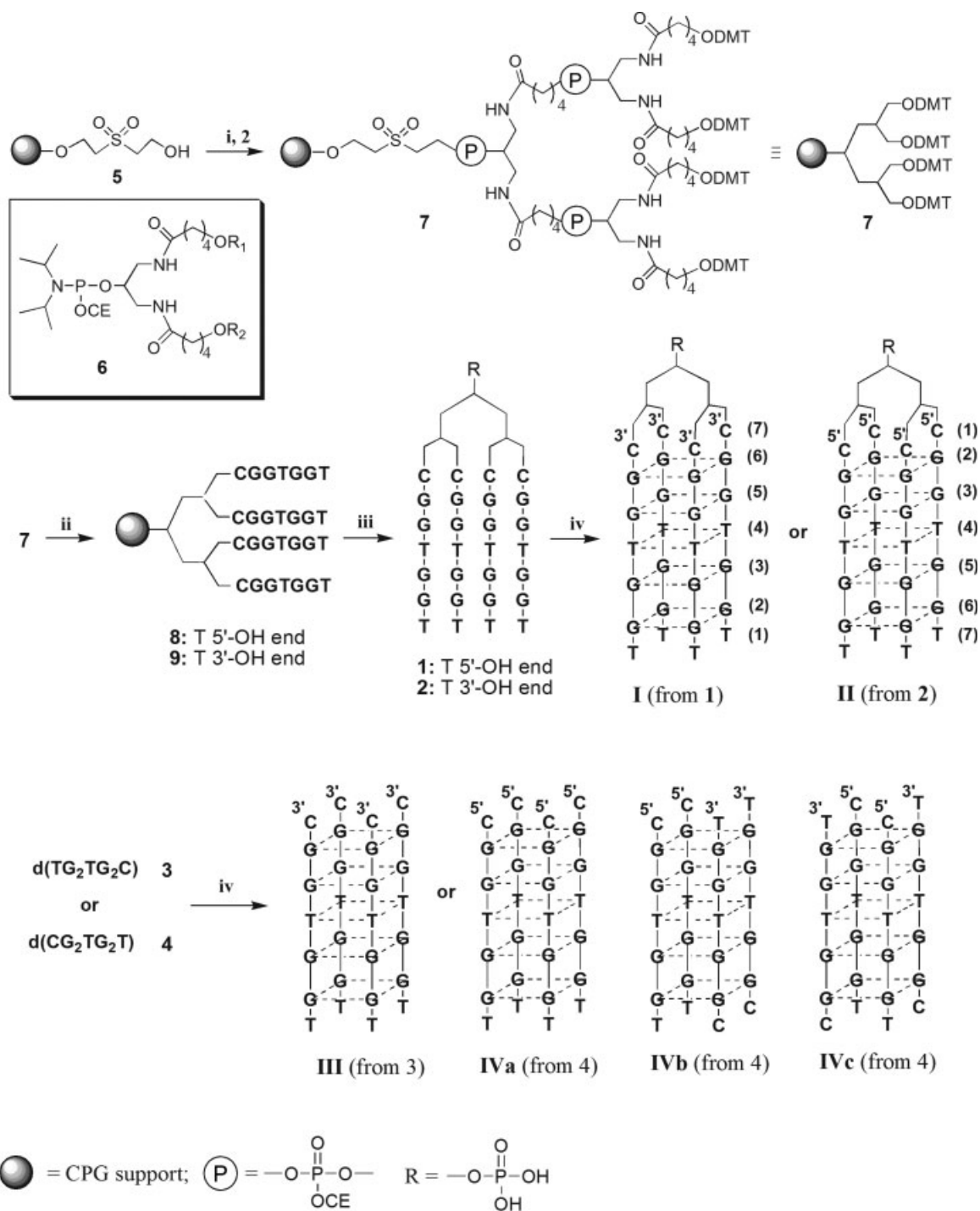


FIGURE 1 G-tetrad and T-tetrad; dR: 2'-deoxyribose.



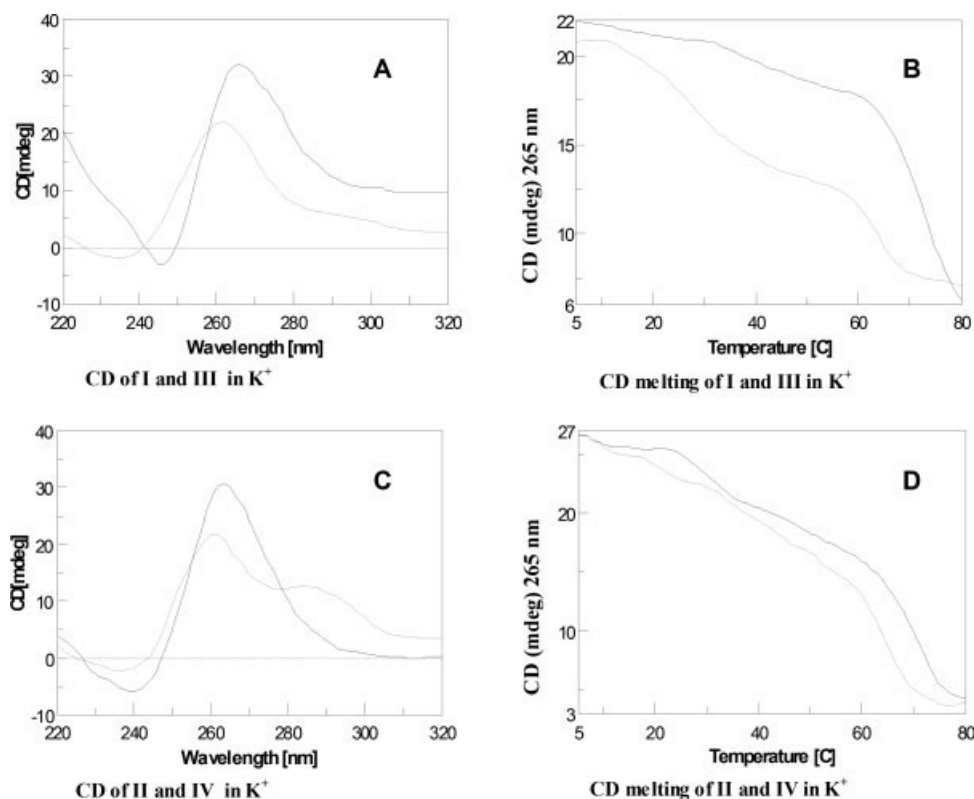
DMT = 4,4'-dimethoxytrityl; CE = 2-cyanoethyl;

**SCHEME 1** (i) two couplings with 2; (ii) ODN synthesis with 3'- or 5'-phosphoramidites; (iii) detachment and deprotection with  $\text{NH}_4\text{OH}$  conc. 32% (7 h,  $55^\circ\text{C}$ ); (iv) HPLC purification and annealing.

their nature and stability in comparison with the tetramolecular complexes **III** and **IV**, we carried out CD studies, including CD thermal denaturation experiments after the annealing procedure. CD spectra

(Figure 2) were performed at  $25^\circ\text{C}$  at a quadruplex concentration of  $2.0 \times 10^{-5} \text{ M}$  in  $\text{K}^+$  buffer. Bunch-ODNs **1**, **2**, as well as natural ODN **3**, showed very similar CD profiles characterized by positive bands at

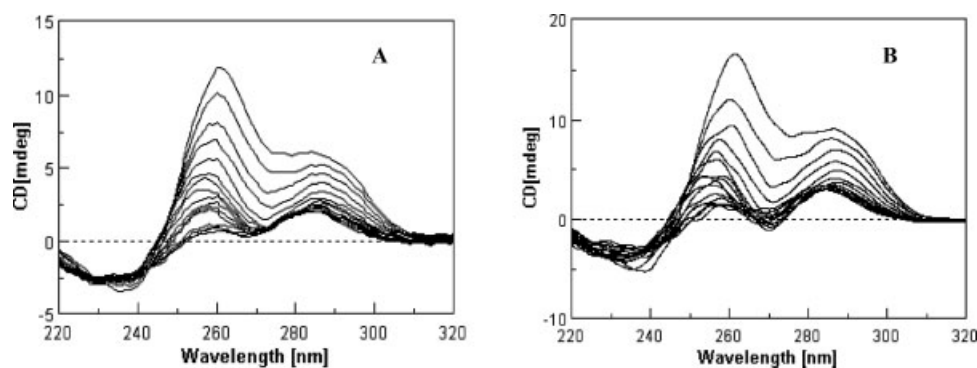




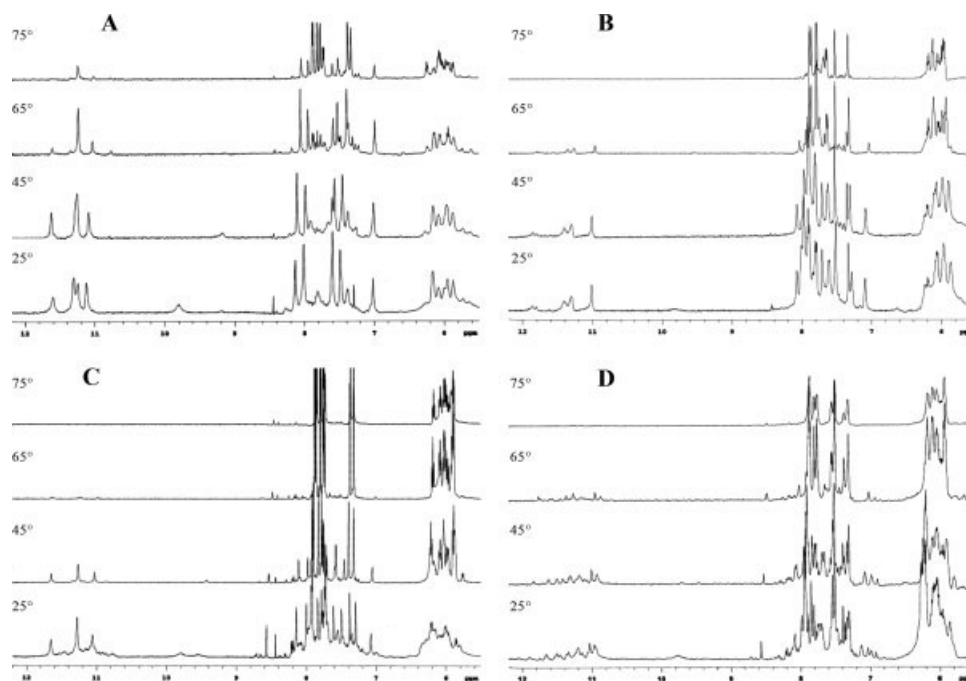
**FIGURE 2** CD spectra at 5°C and melting profiles of quadruplexes **I–II** (solid line) and **III–IV** (dotted line) in K<sup>+</sup> buffer.

266, 264, and 262 nm, respectively, and negative bands at 246, 240, and 238 nm, respectively (Figure 2A,C), which are indicative of parallel stranded G-quadruplex structures<sup>18,19–22</sup> (**I**, **II**, and **III**). However, the CD spectrum of **4** showed a different profile characterized by the presence of two positive bands at 262 and 289 nm and a negative band at 239 nm (Figure 2C). To further investigate the unexpected CD profile showed by **4**, we registered CD spectra in the range 5–80°C (melting process) and 80–5°C (fast annealing process) at 1°C/min scan rate (Figure 3A and 3B,

respectively). In each process, the CD values of the positive bands at 262 and 289 nm varied in a nonconstant ratio. These data suggest that the two positive bands can be attributed to different complexes. Considering that the majority of publications have reported that most antiparallel G-quadruplexes are characterized by a positive band centred at 290–295 nm and a negative band centred around 265 nm,<sup>19–22</sup> it is reasonable to hypothesize that **4** leads to a mixture of quadruplex species probably containing parallel (**IVa**) and antiparallel complexes (**IVb**, **IVc**).



**FIGURE 3** CD spectra of **IV**. (A) Melting process 5–80°C; (B) fast annealing process 80–5°C, at 1°C/min scan rate.



**FIGURE 4** Aromatic and imino proton region of 500-MHz  $^1\text{H}$ -NMR spectra of **I–IV** (A–D, respectively) in 100 mM  $\text{K}^+$  buffer.

### CD Thermal Analysis

CD thermal denaturation experiments were performed in  $\text{K}^+$  buffer monitoring the CD value (mdeg) at 265 nm in the range 5–80°C with 1°C/min heating rate. The melting curve of **I** (Figure 2B) did not show a well-defined change of convexity up to 80°C; however, a drastic reduction of the  $\text{CD}_{265}$  value was observed in the range 60–80°C. The contrary is true in the melting profile of the tetramolecular complex **III** (Figure 2B), since 15–20°C, a significant reduction of the  $\text{CD}_{265}$  value, and two detectable changes of convexity were observed. These data suggest that the helix–coil transition point for complex **I** is at a higher temperature and is more cooperative than that of the corresponding tetramolecular complex **III**. The melting curve of **II** was shown to be very similar to that of **IV** (Figure 2D), both being characterized by the reduction of  $\text{CD}_{265}$  values up to 55–60°C followed by a point of inflection and the complete melting at 80°C, even if a shift toward higher temperature is clearly noticeable for the CD profile of **II**. The whole of the CD melting data suggests that bunch-ODNs **I** and **II** are more stable than their natural counterparts **III** and **IV**, respectively, although a quantitative determination of the relative thermal stability was not performed yet, due to the very slow kinetics of quadruplex denaturation process and considering that, in the case of **II** and **IV**, the latter could be a mixture of different complexes.

### $^1\text{H}$ -NMR Studies on Quadruplexes **I–IV**

Temperature-dependent one-dimensional spectra (25, 45, 65, and 75°C) and NOESY spectra (25°C) were recorded in  $\text{H}_2\text{O}/\text{D}_2\text{O}$  (9:1, v/v) in  $\text{K}^+$  buffer using pulse field gradient WATERGATE for  $\text{H}_2\text{O}$  suppression.<sup>23</sup>

$^1\text{H}$ -NMR spectrum of complex **I** at 25°C (Figure 4A) showed the presence of four well-resolved signals between 11.1 and 11.6 ppm attributable to exchange-protected imino protons involved in Hoogsteen N(1)H/O(6) hydrogen bonds of G-quartets<sup>24,25</sup> also on the basis of the comparison with the corresponding region of the spectrum of **III** that forms a stable parallel-stranded quadruplex already characterized by NMR data.<sup>12</sup> In **I**, the broad signal centred around 9.8 ppm is attributable to H-bonded  $\text{GNH}_2$  protons, while the weak signal at 9.2 ppm is attributable to  $\text{T}_4\text{NH}$  imino protons involved in N(3)H/O(4) hydrogen bonds, thus confirming the formation of a T-tetrad comprised between the  $\text{G}_3$  and  $\text{G}_5$  tetrad planes.  $^1\text{H}$ -NMR spectra at higher temperatures indicated that the quadruplex **I** is preserved at least up to 65°C, whereas the tetramolecular counterpart **III** is almost completely destructured at this temperature, as indicated by the disappearance of the imino proton signals (Figure 4C). Assignment of all aromatic and imino proton signals of **I** was carried out by mean of a NOESY spectrum (25°C; 300 ms mixing time) and comparison of NOEs

with those of **III** reported by Patel and Hosur<sup>12</sup> (data not shown). The pattern of H8/H6–H1'–H8/H6 NOE connectivities for **I** resulted almost completely superimposable to that reported for the corresponding tetramolecular quadruplex **III**, with the exceptions of G<sub>6</sub>H8 signal, upfield shifted from 7.75 to 7.66 ppm, and of C<sub>7</sub>H6 signal downfield shifted from 7.53 to 7.63 ppm. Further, the presence of strong interstrand T<sub>4</sub>NH–T<sub>4</sub>CH<sub>3</sub> and intrastrand T<sub>4</sub>CH<sub>3</sub>–G<sub>5</sub>NH NOEs confirms that **I** adopts a T tetrad containing parallel stranded quadruplex structure. The absence of significant chemical shift drifts for aromatic and imino proton signals of **I** and **III** demonstrates that the “bunch” linker is capable to stabilize the parallel quadruplex structure of **I** without any substantial influence on quadruplex topology.

<sup>1</sup>H-NMR spectrum of compound **II** recorded at 25°C (Figure 2B) supports the CD data about the aptitude of **II** to form a parallel quadruplex structure. In fact, the low field region of <sup>1</sup>H-NMR spectrum of **II** shows the presence of three strong and two weak partially overlapped signals attributable to exchange protected imino protons of G-quartets. Recording <sup>1</sup>H-NMR spectra of **II** at higher temperatures, with 10°C increasing steps, the sharpening and the reduction of intensity of all imino proton signals were observed, followed by their complete disappearance at 75°C. Differently from **I**, no T-tetrad imino signals were observed in the <sup>1</sup>H-NMR spectra of **II**. The low field region of <sup>1</sup>H-NMR spectrum of **IV** recorded at 25°C (Figure 2D) is characterized by the presence of several H-bonded imino signals. The number of imino signals revealed that more than one quadruplex structure must be present in solution. At higher temperatures, with 10°C increasing steps, the disappearance of some signals and the sharpening of the residual ones were observed. Particularly, five major imino signals were detected at 65°C, four of which closely matched those observed for the parallel-stranded quadruplex **II** at the same temperature. This observation supports the hypothesis that **4** folds into more than one quadruplex structures. Considering that CD spectra of **IV** showed a positive shoulder band at 289 nm, it is reasonable to hypothesize that **4** in K<sup>+</sup> buffer could fold into a mixture of quadruplex structures, one of which could be of the antiparallel type. If this were the case, it would be the first occurrence of a tetramolecular antiparallel quadruplex structure.

## CONCLUSIONS

In this article, we present the syntheses and investigations of two bunch-ODNs, **1** and **2**, characterized by

ODN sequences potentially able to form T-tetrad containing quadruplexes. Bunch-ODN **1** is based on the sequence TG<sub>2</sub>TG<sub>2</sub>C, whose natural counterpart represents, to our knowledge, the first example of an experimentally detected T-tetrad. Collected data from NMR and CD melting experiments clearly indicate that the presence of the tetrabranched linker in **1** not only does not hamper the formation of a T-tetrad, but also provides a higher thermal stability to the complex. The capacity of the tetrabranched linker to stabilize a T-tetrad containing quadruplex structure might be particularly interesting in the research area of quadruplex-based aptamers. In fact, the presence of a T-tetrad, thanks to the methyl groups located on the groove surface, could make achievable new hydrophobic interactions capable of improving both the affinity and specificity of a given aptamer toward their target molecule.

The results described in this article confirm the potential of the bunch-ODNs, as a new class of ODN analogues prone to form intramolecular quadruplex structures, to be more stable than the corresponding tetramolecular counterparts when a less stable T-tetrad is present in the complex. So the “bunch” approach to the DNA quadruplex can be considered a useful tool for their *in vivo* applications, furnishing complexes characterized by high stability and predictable structure even at low ODN concentrations. Furthermore, parallel monomolecular bunch-quadruplexes could be used as models in structural studies, being capable to select the formation of only one type of complex, that in some cases is not achievable for tetramolecular complexes.

## EXPERIMENTAL

### Reagents and Equipment

Chemicals and solvents were purchased from Fluka-Sigma-Aldrich. Reagents and phosphoramidites for DNA syntheses were purchased from Glen Research. The ODN were assembled with a PerSeptive Biosystems Expedite DNA synthesizer using phosphoramidite chemistry. HPLC analyses and purifications were performed by a JASCO PU2089 pump system equipped with an ultraviolet (UV) detector 2075 Plus. A Nucleogel SAX (Macherey-Nagel 1000-8/46) and a RP-18 (Thermo, HYPERSIL, 100-C18, 250 × 4.6, 5 μm) columns were used for the purifications. CPG support **5** and phosphoramidite linker **6** were purchased from Glen Research. UV spectra were run by a Jasco V 530 spectrophotometer. CD spectra and thermal denaturation experiments were obtained with a Jasco 715 circular dichroism spectrophotometer equipped with a Jasco 505T temperature controller unit. NMR spectra were recorded with Varian Unity-INOVA 500- and 700-MHz spectrometers. MALDI-

TOF (time of flight) mass spectrometric analysis was performed on a PerSeptive Biosystems voyager-De Pro MALDI mass spectrometer using a picolinic/3-hydroxypicolinic acids mixture as the matrix. The concentration of the ODNs was determined spectrophotometrically at  $\lambda = 260$  nm and at 90°C, using the molar extinction coefficient calculated for the unstacked oligonucleotide using the following extinction coefficient  $\varepsilon = 64100 \text{ cm}^{-1} \text{ M}^{-1}$ , calculated by the nearest neighbor mode.<sup>26</sup>

## Synthesis and Purification of Bunch-Oligonucleotides 1 and 2

Fifty milligrams of support **5** (0.18 meq/g) reacted with phosphoramidite **6** (45 mg/mL in  $\text{CH}_3\text{CN}$ ) following two coupling cycles of standard phosphoramidite chemistry on a 15- $\mu\text{mol}$  scale, yielded support **7** [0.17 meq/g of 4, 4'-dimethoxytrityl (DMT) groups]. After removal of the DMT protecting groups, the ODN chains were assembled, using 3' or 5'-phosphoramidite nucleotide building blocks (45 mg/mL in  $\text{CH}_3\text{CN}$ , 15- $\mu\text{mol}$  scale) to give the polymer bound ODNs **8** and **9**, respectively. The solid support **8** (or **9**) was then treated with 25% aqueous ammonia solution for 7 h at 55°C. The filtered solutions and washings were concentrated under reduced pressure and purified by HPLC on a Nucleogel SAX column eluted with a linear gradient of the following buffers. Buffer A: 20 mM  $\text{KH}_2\text{PO}_4$  aqueous solution pH = 7 containing 20%  $\text{CH}_3\text{CN}$ ; buffer B: 1M KCl, 20 mM  $\text{KH}_2\text{PO}_4$  aqueous solution pH = 7.0 containing 20%  $\text{CH}_3\text{CN}$ ; linear gradient from 0 to 100% B in 30 min, flow rate 1 mL/min. The collected products with retention times 23.0 and 20.1 min for **1** and **2**, respectively, were desalted by gel filtration on a Sephadex G25 column eluted with  $\text{H}_2\text{O}$ /ethanol (9:1, v/v). The purity of the products was confirmed by HPLC analysis on a HYPERSIL 100-C18 column eluted with a linear gradient  $\text{CH}_3\text{CN}$  in triethylammonium bicarbonate (TEAB) buffer (pH 7.0, from 0 to 100% in 60 min, flow 1.0 mL/min), with a result higher than 98% (retention time 13.5 and 14.2 min for **1** and **2**, respectively). The desalifications furnished pure **1** and **2** (68 and 73 OD<sub>260</sub> units, respectively) which, after lyophilization, were characterized by spectroscopic data.

- Bunch-ODN **1**: MALDI-TOF-MS (negative mode) calculated mass: 9070.8; found: 9069.5.
- Bunch-ODN **2**: MALDI-TOF-MS (negative mode) calculated mass: 9070.8; found: 9069.6.

## Preparation of Quadruple Helices (Annealing)

Quadruplexes **I–IV** were formed by dissolving the bunch-ODNs **1–4**, in 10 mM  $\text{K}_2\text{HPO}_4$  100 mM KCl, 0.2 mM EDTA ( $\text{K}^+$  buffer) and annealed by heating to 90°C for 20 min followed by slow cooling to room temperature. The solutions were equilibrated at 5°C for 24 h before performing the experiments.

## CD Experiments

CD spectra of the quadruplexes were registered on a Jasco 715 circular dichroism spectrophotometer in a 0.1-cm path-length cuvette at 20°C. The wavelength was varied from 220 to 320 nm at 5 nm min<sup>-1</sup>. The spectra were recorded with a response of 16 s, at 2.0 nm bandwidth, and normalized by subtraction of the background scan with buffer. CD thermal denaturation experiments were performed monitoring the CD value (mdeg) at 265 nm in the range 5–80°C with 1°C/min heating rate. The CD melting experiments were registered at  $2.0 \times 10^{-5} \text{ M}$  concentration.

## <sup>1</sup>H-NMR Experiments

NMR data were obtained either on a Varian Unity-INOVA 700 MHz or a Varian Unity-INOVA 500 MHz spectrometer equipped with a broad-band inverse probe with a z-field gradient, and processed using Varian VNMR software package. NMR samples of **I–IV** were prepared in  $\text{H}_2\text{O}/\text{D}_2\text{O}$  9:1 at a concentration of single strand of approximately 1.3 mM with a final salt concentration of 100 mM KCl and 10 mM  $\text{K}_2\text{HPO}_4$ . One-dimensional (1D) NMR spectra (500 MHz) were acquired as 16,384 data points with a recycle delay of 1.0 s at temperatures in the range 25–75°C. Data sets were zero filled to 32,768 points and apodized with a squared shifted sine-bell window function. A 300-ms two-dimensional (2D) NOESY spectrum (700 MHz) was acquired at 25°C as a matrix of  $512 \times 2048$  complex points. Relaxation delay was kept to 1.0 s and a squared shifted sine-bell window function was applied in both dimensions prior to Fourier transformation. A pulsed-field gradient WATERGATE sequence<sup>23</sup> was used for  $\text{H}_2\text{O}$  suppression in both 1D and 2D spectra.

## REFERENCES

1. Davis, J. T. *Angew Chem Int Ed* 2004, 43, 668–698.
2. Huppert, J. L.; Balasubramanian, S. *Nucleic Acids Res* 2005, 33, 2908–2916.
3. Todd, A. K.; Johnston, M.; Neidle, S. *Nucleic Acids Res* 2005, 33, 2901–2907.
4. Schaffitzel, C.; Berger, I.; Postberg, J.; Hanes, J.; Lipps, H. J.; Plueckthun, A. *Proc Natl Acad Sci USA* 2001, 98, 8572–8577.
5. Chang, C. C.; Kuo, I. C.; Ling, I. F.; Chen, C. T.; Chen, H. C.; Lou, P. J.; Lin, J. J.; Chang, T. C. *Anal Chem* 2004, 76, 4490–4494.
6. Duquette, M. L.; Handa, P.; Vincent, J. A.; Taylor, A. F.; Maizels, N. *Genes Dev* 2004, 18, 1618–1629.
7. Baldrich, E.; O'Sullivan, C. K. *Anal Biochem* 2005, 341, 194–197.
8. Chou, S. H.; Chin, K. H.; Wang, A. H. J. *Trends Biochem Sci* 2005, 30, 231–234.
9. Patel, P. K.; Koti, A. S. R.; Hosur, R. V. *Nucleic Acids Res* 1999, 27, 3836–3843.
10. Gavathiotis, E.; Searle, M. S. *Org Biomol Chem* 2003, 1, 1650–1656.

11. Searle, M. S.; Williams, H. E. L.; Gallagher, C. T.; Grant, R. J.; Stevens, M. F. G. *Org Biomol Chem* 2004, 2, 810–812.
12. Patel, P. K.; Hosur, R. V. *Nucleic Acids Res* 1999, 27, 2457–2464.
13. Caceres, C.; Wright, G.; Gouyette, C.; Parkinson, G.; Subirana, J. A. *Nucleic Acids Res* 2004, 32, 1097–1102.
14. Patel, P. K.; Bhavesh, N. S.; Hosur, R. V. *NMR Biochem Biophys Res Comm* 2000, 270, 967–971.
15. For a recent example see: Webba da Silva, M. *Biochemistry* 2005, 44, 3754–3764.
16. Chen, J.; Zhang, L. R.; Min, J. M.; Zhang, L. H. *Nucleic Acids Res* 2002, 30, 3005–3014.
17. Virgilio, A.; Esposito, V.; Randazzo, A.; Mayol, L.; Galeone, A. *Bioorg Med Chem* 2005, 13, 1037–1044.
18. Oliviero, G.; Borbone, N.; Galeone, A.; Varra, M.; Piccialli, G.; Mayol, L. *Tetrahedron Lett* 2004, 45, 4869–4872.
19. Petraccone, L.; Erra, E.; Esposito, V.; Randazzo, A.; Mayol, L.; Nasti, L.; Barone, G.; Giancola, C. *Biochemistry* 2004, 43, 4877–4884.
20. Jin, R.; Gaffney, B. L.; Wang, C.; Jones, R. A.; Breslauer, K. J. *Proc Natl Acad Sci USA* 1992, 89, 8832–8836.
21. Dapić, V.; Abdomerović, V.; Marrington, R.; Pederby, J.; Rodger, A.; Trent, J. O.; Bates, P. *Nucleic Acids Res* 2003, 31, 2097–2170.
22. Lu, M.; Guo, Q.; Kallenbach, N. R. *Biochemistry* 1993, 32, 598–601.
23. Piotto, M.; Saudek, V.; Sklenar, V. J. *Biomol NMR* 1992, 2, 661–665.
24. Feigon, J.; Koshlap, K. M.; Smith, F. W. *Methods Enzymol* 1995, 261, 225–255.
25. Feigon, J. *Encyclopedia of Nuclear Magnetic Resonance*; John Wiley & Sons: West Sussex, 1996; pp 1726–1731.
26. Breslauer, K. J.; Frank, R.; Blöcker, H.; Marky, L. A. *Proc Natl Acad Sci USA* 1986, 83, 3746–3750.

*Reviewing Editor: Kenneth Breslauer*



# Synthesis and Characterization of Monomolecular DNA G-Quadruplexes Formed by Tetra-End-Linked Oligonucleotides

Giorgia Oliviero,<sup>†</sup> Jussara Amato,<sup>†</sup> Nicola Borbone,<sup>†</sup> Aldo Galeone,<sup>†</sup> Luigi Petraccone,<sup>‡</sup> Michela Varra,<sup>†</sup> Gennaro Piccialli,<sup>\*,†</sup> and Luciano Mayol<sup>†</sup>

Dipartimento di Chimica delle Sostanze Naturali, Università degli Studi di Napoli "Federico II", Via D. Montesano 49, I-80131 Napoli, Italy, and Dipartimento di Chimica, Università degli Studi di Napoli "Federico II", Via Cintia, I-80126 Napoli, Italy. Received January 17, 2006; Revised Manuscript Received May 18, 2006

Guanine-rich DNA sequences are widely dispersed in the eukaryotic genome and are abundant in regions with relevant biological significance. They can form quadruplex structures stabilized by guanine quartets. These structures differ for number and strand polarity, loop composition, and conformation. We report here the syntheses and the structural studies of a set of interconnected d(TG<sub>4</sub>T) fragments which are tethered, with different orientations, to a tetra-end-linker in an attempt to force the formation of specific four-stranded DNA quadruplex structures. Two synthetic strategies have been used to obtain oligodeoxyribonucleotide (ODN) strands linked with their 3'- or 5'-ends to each of the four arms of the linker. The first approach allowed the synthesis of tetra-end-linked ODN (TEL-ODN) containing the four ODN strands with a parallel orientation, while the latter synthetic pathway led to the synthesis of TEL-ODNs each containing antiparallel ODN pairs. The influence of the linker at 3'- or 5'-ODN, on the quadruplex typology and stability, in the presence of sodium or potassium ions, has been investigated by circular dichroism (CD), CD thermal denaturation, <sup>1</sup>H NMR experiments at variable temperature, and molecular modeling. All synthesized TEL-ODNs formed parallel G-quadruplex structures. Particularly, the TEL-ODN containing all parallel ODN tracts formed very stable parallel G-quadruplex complexes, whereas the TEL-ODNs containing antiparallel ODN pairs led to relatively less stable parallel G-quadruplexes. The molecular modeling data suggested that the above antiparallel TEL-ODNs can adopt parallel G-quadruplex structures thanks to a considerable folding of the tetra-end-linker around the whole quadruplex scaffold.

## INTRODUCTION

DNA quadruple helices (quadruplexes) based on guanine (G)<sup>1</sup> quartets have recently emerged as biologically important structures (1–4). G-rich sequences, which are found primarily in telomeric region at 3'-end of chromosomes, are able to fold into G-quadruplex structures. It has been suggested that quadruplex formation could inhibit the telomere maintenance provided by telomerase activity thus affecting the lifespan of the cells of a number of cancer types characterized by a high level of telomerase expression (5–8). Furthermore, studies on the oncogene *c-MYC*, involved in a number of malignancies, have demonstrated the existence of G-quadruplex structures in vivo which can be considered useful targets to repress *c-MYC* transcription (9–11). G-quadruplex structures are present in the scaffold of several aptamers provided with useful biological properties (12–16), as well as within the insulin gene-linked polymorphic region of the human insulin gene, in vivo, where they may play a role in the expression of the insulin gene (17). G-quadruplexes, have been also put forth as DNA structures capable of transferring electrons through the G-quadruplex  $\pi$ -stack and forming nanowires for molecular nanoelectronics (18–20). Within this framework, a number of studies have been

undertaken to investigate the DNA G-quadruplexes and, particularly, to define the factors that can stabilize them as well as proposing synthetic ODN analogues capable of forming novel quadruplex typologies.

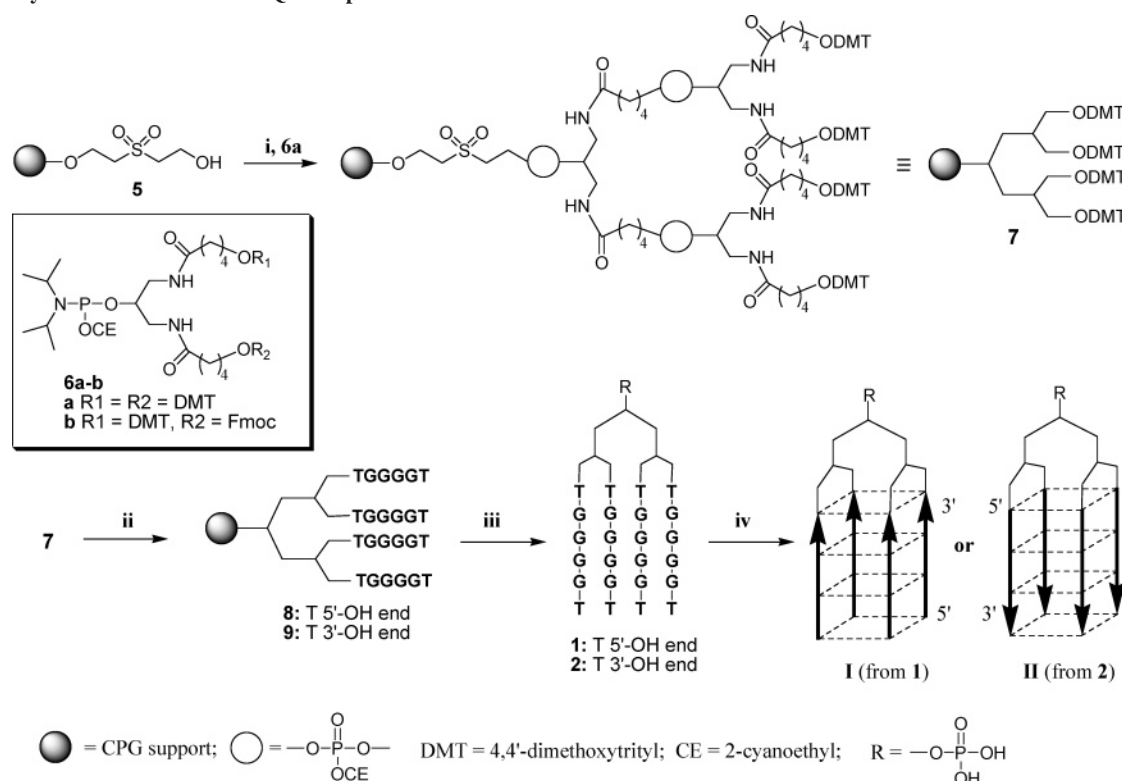
G-quadruplexes can be classified based upon the number of strands involved in the structure, their relative orientation and the glycosidic conformation of guanosine residues (anti or syn). The stability of G-quadruplexes mainly depends on the number of G-tetrad planes, the molecularity of the complex, and the nature of cations which accommodate themselves in the central channel of the quadruplex lowering the energy of the complex (1–3). Normally, the molecularity of the complex induces well defined patterns in the relative ODN strand orientation. The monomolecular complexes largely involved in biomolecular events (1–4) are prevalently of antiparallel type even if examples of parallel monomolecular complexes have been discovered (21–24). Tetramolecular complexes, which have recently displayed interesting biological properties (25–27), always show the four strands in a parallel orientation. However, the development of therapeutic applications for such ODN complexes has been so far prevented by their unfavorable kinetic and thermodynamic parameters. It is a fact that the in vitro formation of a tetramolecular quadruplex is very slow and may require high ODN and salt concentrations (28). On the contrary, monomolecular quadruplexes are mostly characterized by high stability and fast kinetic formation (1–3). On these bases, the construction of new and stable monomolecular quadruplexes possessing unusual strand orientation could expand their employment in both the study and control of important biological processes as well as in DNA-based nanotechnologies. Several studies have investigated the effect of loops on quadruplex stability and typology (29–35). With regard to

\* Corresponding author. E-mail: piccialli@unina.it. Voice: +39-081-678541. Fax: +39-081-678552.

<sup>†</sup> Dipartimento di Chimica delle Sostanze Naturali.

<sup>‡</sup> Dipartimento di Chimica.

<sup>1</sup> Abbreviations: ODN, oligodeoxyribonucleotide; TEL, tetra-end-linked; CD, circular dichroism; G, guanine; Fmoc, fluorenylmethoxycarbonyl; DMT, 4,4'-dimethoxytrityl; NMI, 1-methylimidazole; MSNT, 1-(2-mesitylenesulfonyl)-3-nitro-1*H*-1,2,4-triazole; DCA, dichloroacetic acid; TEAB, triethylammonium bicarbonate; CPG, controlled pore glass; TBE, Tris-borate electrophoresis.

Scheme 1. Synthetic Procedures for Quadruplexes I and II<sup>a</sup>

<sup>a</sup> i: two couplings with **6a**; ii: ODN synthesis with 3' or 5'-phosphoramidites; iii: detachment and deprotection with  $\text{NH}_4\text{OH}$  concd 32% (7 h, 55 °C); iv: HPLC purification and annealing.

nucleotidic loops, the base composition, the length, and the orientation of the phosphodiester linkage have already been examined (29–34). In addition, mono- or bimolecular quadruplexes formed by circular ODNs have been shown to display noteworthy thermal and enzymatic degradation stability (35). In another reported approach aimed at studying the effect of the connecting loops on the stability and typology of the monomolecular quadruplexes, non nucleotidic linkers have been inserted between d(GGG) tracts into a single ODN strand (30). In this case both the kinetics of the quadruplex formation and the complex stabilities were found to be affected by the nature of the linker.

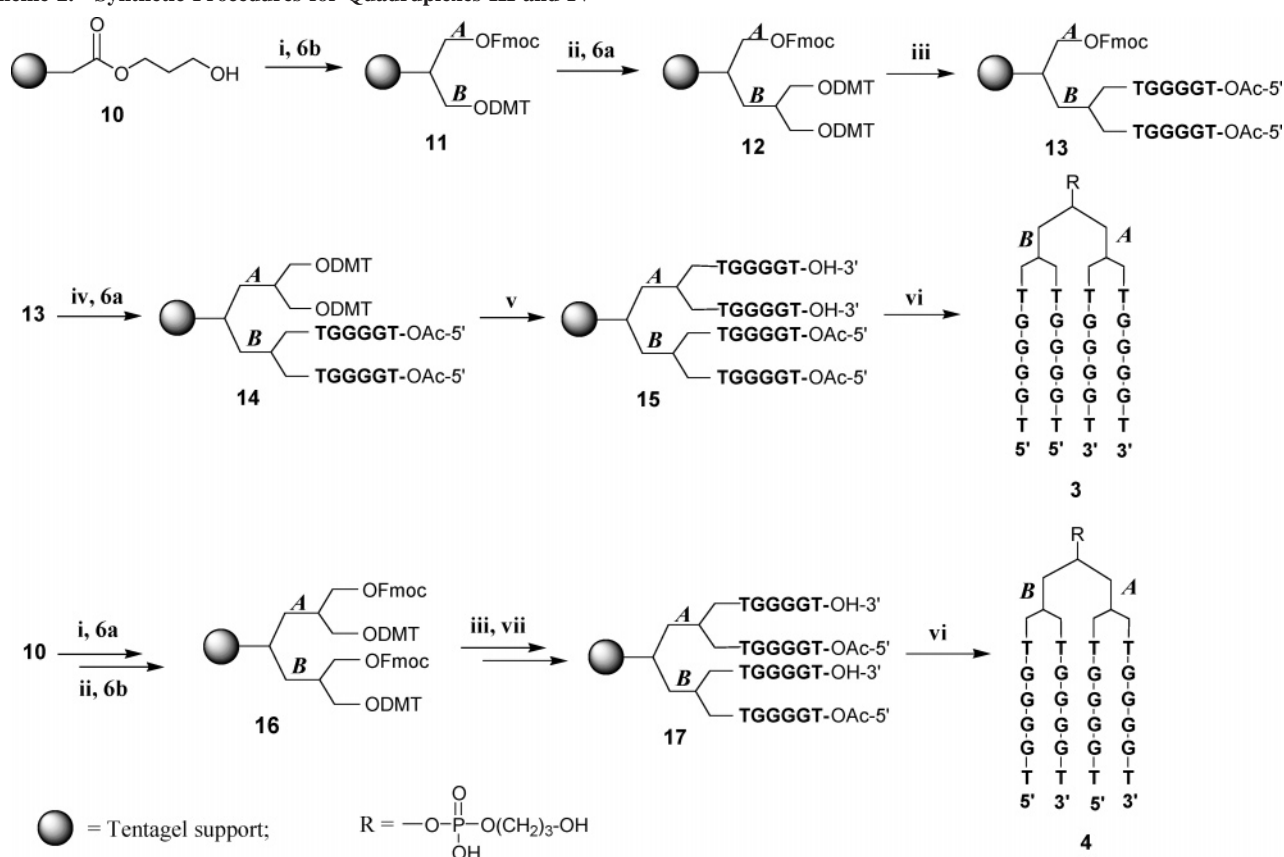
Starting from these observations, we recently proposed the synthesis of a new modified ODN (TEL-ODN, **1**, Scheme 1) capable of forming the monomolecular parallel G-quadruplex **I** (Scheme 1) whose thermal stability has been proved to be significantly higher than that observed for its tetramolecular counterpart  $[\text{d}(\text{TG}_4\text{T})]_4$  (36). The main structural feature of this molecule is the presence of four ODN strands whose 3'-ends are attached to a non nucleotidic tetra-end-linker. Furthermore, TEL-ODNs were also shown to be a useful tool to study quadruplexes containing less stable quartets such as T-tetrad (37).

We wish to report here our study on the influence of the tetra-end-linker placement on the typology and stability of this kind of quadruplexes. Two synthetic strategies have been used to obtain d(TG<sub>4</sub>T) ODN strands which were linked with their 3' or 5'-ends to each of the four arms of the linker. The first approach allowed the synthesis of the TEL-ODNs containing the four ODN strands with a parallel orientation (**1**, **2**, Scheme 1) while the latter synthetic pathway led to the synthesis of TEL-ODNs each containing antiparallel ODN pairs (**3**, **4**, Scheme 2). CD, CD melting, NMR spectroscopy, and molecular modeling have been used to investigate the solution structures of the resulting TEL-quadruplexes **I–IV** (Schemes 1 and 3) formed in the presence of sodium or potassium ions.

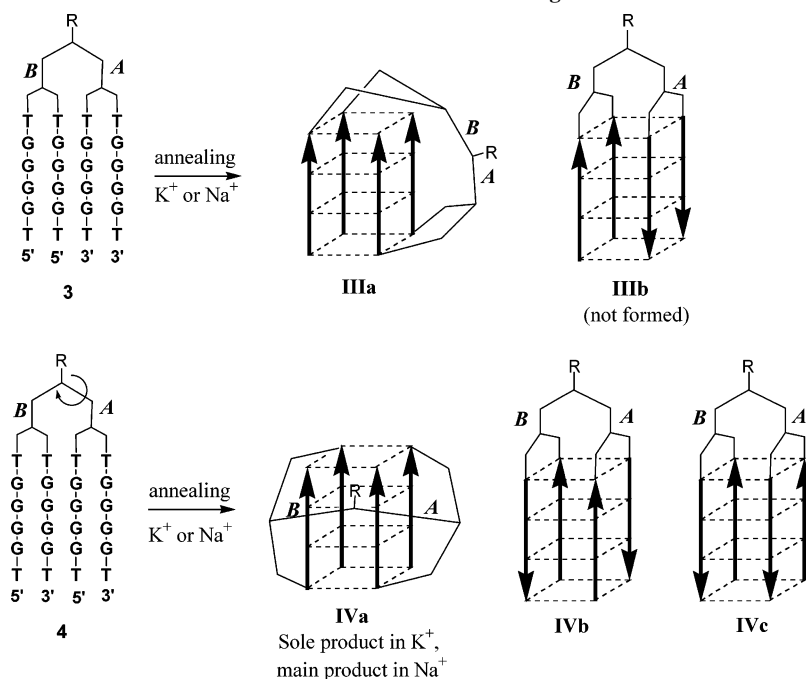
## EXPERIMENTAL PROCEDURES

**Reagents and Equipment.** Chemicals and solvents were purchased from Fluka-Sigma-Aldrich. Reagents and phosphoramidites for DNA syntheses were purchased from Glen Research. ODN syntheses were performed on a PerSeptive Biosystems Expedite. HPLC analyses and purifications were performed by a JASCO PU2089 pumps equipped with UV detector 2075 Plus using a Merck Hibar (5  $\mu\text{m}$ , 250–10) column. CPG support **5** and linkers **6a,b** were purchased from Glen Research. Tentagel carboxy resin, used as starting material for preparation of solid support **10**, was purchased from Novabiochem. Preparation of **10** was carried out in a short column (10 cm length, 1 cm i.d.) equipped with a sintered glass filter, a stopcock and a cap. The ODNs were assembled by a PerSeptive Biosystems Expedite DNA synthesizer using phosphoramidite chemistry. UV spectra were run with a Jasco V 530 spectrophotometer. CD spectra and thermal denaturation experiments were obtained with a Jasco 715 circular dichroism spectrophotometer equipped with a JASCO 505T temperature controller unit. NMR spectra were recorded on a Varian Unity Inova 500 MHz spectrometer. MALDI-TOF mass spectrometric analyses were performed on a PerSeptive Biosystems voyager-De Pro MALDI mass spectrometer using a picolinic/3-hydroxypicolinic acids mixture as the matrix. The ODN concentration was determined spectrophotometrically at  $\lambda = 260 \text{ nm}$  and 90 °C, using the molar extinction coefficient  $\epsilon = 57800 \text{ cm}^{-1} \text{ M}^{-1}$  calculated for the unstacked oligonucleotide by the nearest neighbor mode (38).

**Synthesis and Purification of TEL-ODNs 1 and 2.** 50 mg of support **5** (0.048 mequiv/g) was used for each synthesis in the automated DNA synthesizer following standard phosphoramidite chemistry, using 45 mg/mL of solution of phosphoramidite **6a** (two coupling cycles) and 3'- or 5'-phosphoramidite nucleotide building block (six cycles) in anhydrous  $\text{CH}_3\text{CN}$ , thus obtaining the polymer bound ODN **8** and **9**, respectively. The coupling efficiency with nucleotide units was consistently

Scheme 2. Synthetic Procedures for Quadruplexes III and IV<sup>a</sup>

<sup>a</sup> i: coupling with linker **6**; ii: DMT removal and coupling with linker **6**; iii: ODN synthesis with 3'-phosphoramidites; iv: Fmoc removal and coupling with linker **6a**; v: ODN synthesis with 5'-phosphoramidites; vi: detachment and deprotection with  $NH_4OH$  concd 32% (7 h, 55 °C) and HPLC purification; vii: Fmoc removal and ODN synthesis with 5'-phosphoramidites;

Scheme 3. Quadruplexes III and IV from TEL-ODNs 3 and 4 after the Annealing Procedure in  $K^+$  or  $Na^+$  Buffer

higher than 98% (by DMT spectrophotometric measurements). The solid support **8** (or **9**) was then treated with 25% aq ammonia solution for 7 h at 55 °C. The filtered solution and washings were concentrated under reduced pressure and purified by HPLC on a RP18 column eluted with a linear gradient  $CH_3CN$  in TEAB buffer (pH 7.0, from 0 to 100% in 60 min, flow 1.0 mL/min), retention time 26.1 and 27.0 min for **1** and

**2**, respectively. The collected peaks furnished pure **1** and **2** (72 and 77 OD<sub>260</sub> units, respectively) which, after lyophilization, were characterized by spectroscopic data.

TEL-ODN **2**: MALDI TOF-MS (negative mode): calculated mass (8763.7); found: 8762.6.

**Preparation of Support 10.** 500 mg of Tentagel carboxy resin (0.28 mequiv/g of COOH) suspended in  $CH_2Cl_2$  were left



in contact with a  $\text{CH}_2\text{Cl}_2$  solution (7.0 mL) containing 1,3-dihydroxypropane (0.56 mmol, 0.36 mL), NMI (0.56 mmol, 0.40 mL) and MSNT (0.056 mmol, 1.48 g) and the mixture was shaken for 36 h at room temperature. The obtained hydroxy-support **10** was exhaustively washed with  $\text{CH}_2\text{Cl}_2$  and then dried under reduced pressure. The amount of hydroxyl groups, determined by DMT measurements after the successive coupling step with 3'-phosphoramidite nucleotide units, resulted to be in the range 0.18–0.20 mequiv/g.

**Synthesis and Purification of TEL-ODNs 3 and 4.** 50 mg of support **10** (0.18 mequiv/g) reacted, in the automated DNA synthesizer, with phosphoramidite **6b** (45 mg/mL in  $\text{CH}_3\text{CN}$ ) following standard phosphoramidite chemistry at 15- $\mu\text{mol}$  scale, yielded support **11** (0.17 mequiv/g of DMT groups). After removal of the DMT group by DCA/standard treatment, the second coupling cycle was performed in the same manner using phosphoramidite **6a**, thus obtaining support **12** (0.30 mequiv/g of DMT groups). **12** was then subjected to six coupling cycles using 3'-phosphoramidite nucleotide building blocks (45 mg/mL in  $\text{CH}_3\text{CN}$ , 15- $\mu\text{mol}$  scale), followed by final DMT removal and 5'-OH capping with  $\text{Ac}_2\text{O}$ , to yield ODN-functionalized support **13**. The removal of Fmoc groups was achieved by treatment with piperidine/DMF solution (2:8 v/v, 30 min RT). The resulting support was then reacted with phosphoramidite **6a**, thus obtaining **14** (0.27 mequiv/g of DMT groups). On branches A of **14** the second ODN domain was assembled, performing six coupling cycles with the 5'-phosphoramidite nucleotide building block as described for support **12**, yielding the polymer bound ODN **15**.

For the synthesis of **4**, support **10** (50 mg) was functionalized with **6a** and then **6b**, as described before, yielding the tetra-branched support **16** (0.28 mequiv/g of DMT groups). After removal of the DMT protecting groups, the first two ODN chains were assembled, using 3'-phosphoramidite nucleotide building blocks. After capping of the terminal 5'-OH functions by  $\text{Ac}_2\text{O}$  treatment, the Fmoc groups were removed as previously described, and the successive two ODN chains were assembled using 5'-phosphoramidite nucleotide building blocks, thus obtaining the polymer bound ODN **17**.

TEL-ODNs were detached from the supports **15** and **17**, deprotected, and purified as described for **1** and **2**, thus obtaining **3** and **4** (retention time 25.8 and 25.2 min, respectively). After lyophilization, the final pure products **3** and **4** (68 and 62 OD<sub>260</sub> units, respectively) were characterized by spectroscopic data.

**TEL-ODN 3:** MALDI TOF-MS (negative mode): calculated mass (8821.7); found: 8820.5.

**TEL-ODN 4:** MALDI TOF-MS (negative mode): calculated mass (8821.7); found: 8820.2

**Preparation of Quadruple Helices (Annealing).** Quadruplexes **I–IV** were formed by dissolving the TEL-ODNs **1–4**, in the appropriate buffer and annealed by heating to 90 °C for 20 min followed by slow cooling to room temperature. The solutions were equilibrated at 10 °C for 24 h before performing the experiments. The buffers used were 10 mM  $\text{Na}_2\text{HPO}_4$ , 70 mM NaCl, 0.2 mM EDTA ( $\text{Na}^+$  buffer) or 10 mM  $\text{K}_2\text{HPO}_4$  70 mM KCl, and 0.2 mM EDTA ( $\text{K}^+$  buffer).

**Native Gel Electrophoreses.** Native gel electrophoreses were run on 20% nondenaturing polyacrylamide gels in 0.6x TBE buffer, pH 7.0, with 50 mM NaCl. ODN solutions, with a final concentration of 300  $\mu\text{M}$  in quadruplex, were prepared in 50 mM Tris-borate pH 7.0, 80 mM NaCl, 1 mM spermine, and 3 mM  $\text{MgCl}_2$  and annealed as above-described. At 2  $\mu\text{L}$  of the ODN solution (300  $\mu\text{M}$  in quadruplex) was then added 1  $\mu\text{L}$  of 30% glycerol and the mixture charged on gel. The gels were run at room temperature at constant voltage (100 V) for 2.5 h. The bands were visualized by UV shadowing.

**CD Experiments.** CD spectra of the quadruplexes were registered on a Jasco 715 circular dichroism spectrophotometer in a 0.1 cm path length cuvette at 25 and 90 °C at  $1.0 \times 10^{-5}$  and  $5.0 \times 10^{-5}$  M quadruplex concentrations. The wavelength was varied from 220 to 320 nm at 5 nm  $\text{min}^{-1}$ . The spectra were recorded with a response of 16 s, at 2.0 nm bandwidth and normalized by subtraction of the background scan with buffer. CD thermal denaturation–renaturation experiments were performed monitoring the CD value (mdeg) at 264 nm in the range 25–90 °C with 1 °C/min heating rate.

**$^1\text{H}$  NMR Experiments.**  $^1\text{H}$  NMR data were collected on a Varian Unity INOVA 500 spectrometer equipped with a broadband inverse probe with z-field gradient and processed using Varian VNMR software package. 1D NMR spectra were acquired as 16384 data points with a recycle delay of 1.0 s at temperatures in the range 25–85 °C. Data sets were zero filled to 32768 points prior to Fourier transformation and apodized with a shifted sinebell squared window function. Pulsed-field gradient WATERGATE (39) sequence was used for  $\text{H}_2\text{O}$  suppression. All NMR samples **I–IV** were prepared in  $\text{H}_2\text{O}/\text{D}_2\text{O}$  (9:1, v/v) at 0.5 mM of quadruplex concentration with a final salt concentration of 100 mM KCl and 10 mM  $\text{K}_2\text{HPO}_4$  or 100 mM NaCl and 10 mM  $\text{Na}_2\text{HPO}_4$  for  $\text{K}^+$  or  $\text{Na}^+$  buffer, respectively.

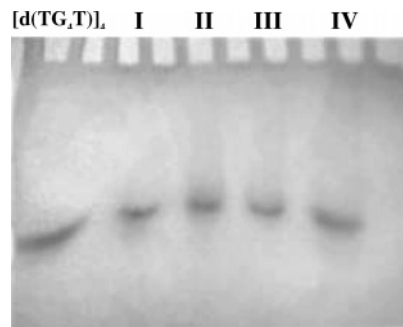
**Molecular Modeling.** The conformational features of the quadruplexes **III** and **IV** have been explored by means of molecular modeling study. All the calculations were performed on a personal computer running the HyperChem 7.5 suite of programs. The AMBER force field based on AMBER 99 parameter set was used (40). The initial coordinates for the starting model of  $[\text{d}(\text{TGGGGT})]_4$  quadruplex were taken from the NMR solution structure of the  $[\text{d}(\text{TTGGGGT})]_4$  quadruplex (Protein Data Bank entry number 139D), choosing randomly one of the four available structures. The initial  $[\text{d}(\text{TGGGGT})]_4$  quadruplex models were built by deletion of the 5'-end thymidine residue in each of the four TTGGGGT strands. The complete structures of **III** and **IV** were then built using the HyperChem 7.5 building tool. Partial charges for each of the tetra-end-linker atoms were assigned using the Gasteiger-Marsili algorithm (41) implemented in the QSAR module of HyperChem suite. The resulting coordinates of tetra-end-linker's atoms were energy-minimized in a vacuum, keeping all DNA coordinates frozen (500 cycles of the steepest descent method). Three sodium ions were manually positioned in the central channel, equidistant from adjacent G-tetrads to allow octahedral coordination with G carbonyl oxygen O6 atoms. The two complexes were neutralized by addition of further 26 sodium ions placed in the most negative positions of the systems. Each quadruplex was then solvated by a periodic TIP3P water box (5833 water molecules) (42), which extended to a distance of 10 Å from each atom. The minimal distance between water and system atoms was set to 2.3 Å. The whole systems were initially subjected to energy minimization followed by 50 ps of molecular dynamics at 300 K, and then the systems were slowly cooled at 10 K. The obtained structures were further minimized using 1000 cycles of the steepest descent method followed by conjugate gradient method until convergence to a rms gradient of 0.1 kcal  $\text{mol}^{-1}$  Å $^{-1}$ .

## RESULTS AND DISCUSSION

**Synthesis and Purification of TEL-ODNs 1–4.** In the synthesis of the TEL-ODNs **1–4** we used two hydroxy-functionalized solid supports (**5** and **10**) and the bifunctional linkers **6a,b**. Support **5** (0.048 mequiv/g), by way of two coupling cycles with **6a** by an automatic DNA synthesizer, was turned into support **7** bearing a symmetrical tetra-end-linker with four protected primary alcoholic functions prone to ODN-chain

assembly (0.085 mequiv/g of OH functions, by DMT spectrophotometric test). Support **7** was used to synthesize the TEL-ODNs **1** and **2** in which the four ODN strands were attached to the tetra-end-linker by the same end (3'- and 5'-end, respectively). ODN chains were assembled on **7** by using 3'- or 5'-phosphoramidite nucleotide building blocks, thus obtaining the polymer bound TEL-ODNs **8** and **9**, respectively. In both syntheses coupling yields were higher than 98% per cycle, thus reducing the amounts of TEL-ODNs having one or more truncated chains. The presence of a  $\beta$ -sulfone phosphodiester function connecting the linker to the solid support, allows the release of the TEL-ODN via a  $\beta$ -elimination under basic treatment. In this way supports **8** and **9**, treated with concentrated  $\text{NH}_4\text{OH}$  (7 h, 55 °C), afforded crude TEL-ODNs **1** and **2**, which after HPLC purification and the annealing procedure, furnished the quadruplex **I** and **II**, respectively.

As a further step of this work we synthesized the TEL-ODNs **3** and **4** (Scheme 2). In **3** a pair of strands was linked to the primary linker branch A through the 5'-ends whereas the second pair is linked to the primary branch B through the 3'-ends. Instead **4** is characterized by both branches (A and B) carrying two strands linked through opposite ends. For the preparation of **3** and **4**, we adopted two synthetic strategies both based upon bifunctional linker **6b** in which the two alcoholic functions, thanks to the orthogonal protection with Fmoc and DMT groups, allowed the sequential synthesis of the ODN chains with opposite polarity. **10** was used as solid support since, when **5** was used, we observed the release of ODN material from the support during the Fmoc deprotection step, due to the concomitant  $\beta$ -elimination at the sulfone group. Therefore, we prepared the hydroxy-functionalized support **10** bearing an ester function, stable under Fmoc deprotection conditions but cleavable by ammonia treatment, to release the assembled TEL-ODN. Support **10** (0.18–0.20 mequiv/g of OH groups), prepared by reacting a carboxy-functionalized Tentagel resin (0.28 mequiv/g of COOH groups) with 1,3-dihydroxypropane in the presence of MSNT and NMI (43), was used in two synthetic pathways leading to the TEL-ODNs **3** and **4**. Support **10**, by reaction with the phosphoramidite linker **6b** yielded the asymmetrical protected bifunctional support **11** (0.17 mequiv/g by DMT test) which, after removal of DMT protection from the branch B and successive reaction with linker **6a** furnished the three-functionalized support **12**. Six coupling cycles with 3'-phosphoramidite units, performed as described for **7**, and a final capping step of the 5'-OH ODN ends, gave **13**. Then the 5'-Fmoc group was removed from the branch A by piperidine treatment, and the successive coupling with **6a** yielded **14**. On **14** the second ODN domain was assembled using 5'-phosphoramidite units, thus obtaining polymer bound TEL-ODN **15**. The synthetic pathway to obtain **4** started with reaction of **10** with **6a** and subsequently **6b** as described above, thus obtaining the tetra-branched support **16** (0.28 mequiv/g by DMT). The removal of DMT groups, followed by ODN synthesis with 3'-phosphoramidite units, allowed the assembly of the first pair of ODN chains. After Fmoc deprotection, the remaining ODN pair having opposite polarity was then assembled using 5'-phosphoramidite units, thus obtaining **17**. Detachment from the solid support and complete deprotection of the TEL-ODNs **3** and **4** were achieved treating supports **15** and **17** with concentrated  $\text{NH}_4\text{OH}$  as described for **1** and **2**. Analyses and purifications of crude products **1–4** were carried out by HPLC using an RP18 column. The purity of the products was checked by PAGE, and the structures were confirmed by  $^1\text{H}$  NMR and MALDI-MS data. The quadruple helices were obtained by dissolving the samples **1–4** in  $\text{Na}^+$  or  $\text{K}^+$  buffers and by the annealing procedure to ensure the correct formation of quadruplex structures.

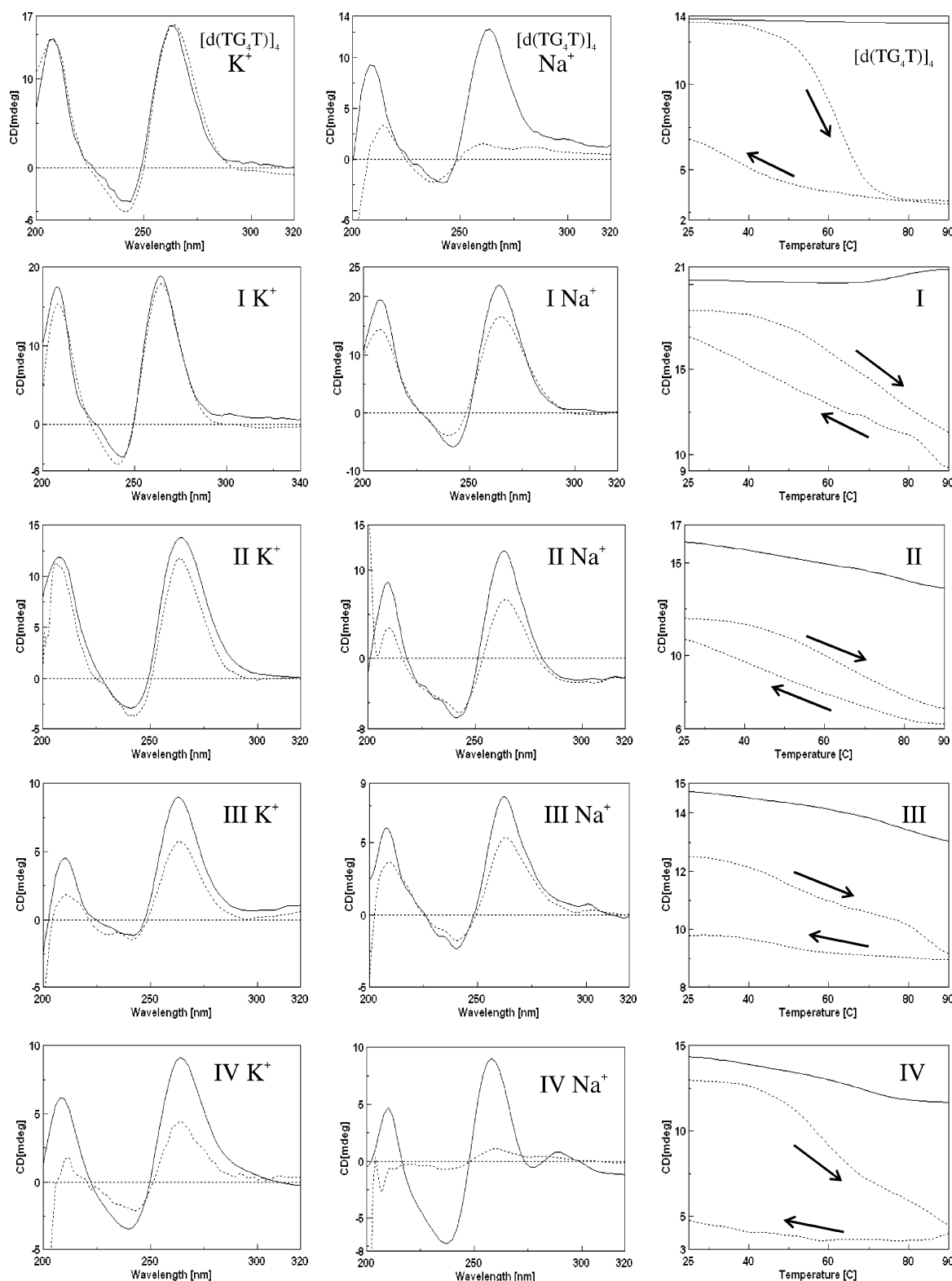


**Figure 1.** Nondenaturing gel analysis of **I–IV** and  $[\text{d}(\text{TG}_4\text{T})]_4$  (20% polyacrylamide), Tris-borate (TBE), pH 7.0. The samples were annealed in  $\text{Na}^+$  buffer. The bands were visualized by UV shadowing.

**Native Gel Electrophoresis.** The purity of the products **1–4** and the molecularity of the complexes obtained were investigated by performing a nondenaturing polyacrylamide gel electrophoresis (Figure 1) in comparison to the tetramolecular quadruplex  $[\text{d}(\text{TG}_4\text{T})]_4$ . The samples (300  $\mu\text{M}$  quadruplex) were annealed in  $\text{Na}^+$  buffer, and the electrophoretic analysis was carried out using a 20% polyacrylamide gel in TBE buffer. TEL-quadruplexes **I–IV** and  $[\text{d}(\text{TG}_4\text{T})]_4$  migrated as single bands demonstrating quite similar mobilities. This result, besides indicating the purities of **I–IV**, strongly suggests that they form monomeric species at this concentration.

**CD and CD Thermal Analysis.** To investigate whether TEL-ODNs **2–4** adopt quadruplex structures and, if such is the case, determine their nature and stability, we carried out CD studies, including CD thermal denaturation experiments, which were compared with those already reported for **1** (36) and their tetramolecular counterpart  $[\text{d}(\text{TG}_4\text{T})]_4$  (27, 38). The ODN samples were analyzed in  $\text{Na}^+$  or  $\text{K}^+$  buffers after the annealing procedure. CD spectra (Figure 2) were performed at 25 and 90 °C at a quadruplex concentration of  $1.0 \times 10^{-5}$  and  $5.0 \times 10^{-5}$  M. TEL-ODNs **1**, **2** and  $[\text{d}(\text{TG}_4\text{T})]_4$  showed almost identical CD profiles in  $\text{K}^+$  buffer at 25 °C, characterized by positive and negative bands centered around 264 and 244 nm, respectively, which are indicative of a parallel stranded G-quadruplex structure (36, 44–47). Upon the basis of these findings, it is reasonable to hypothesize that the TEL-ODN **2**, bearing the tetra-end-linker at the 5'-end of quadruplex, adopts the parallel G-quadruplex structure **II**. The data also suggest that **3** and **4** adopt parallel G-quadruplex structures which, considering the mutual antiparallel orientation of the ODN pairs, can be explained only by hypothesizing that a considerable folding of the linker around the whole G-quadruplex takes place. The parallel structures adopted by **3** and **4** were substantiated by molecular modeling studies (see below) that clearly indicate that these conformations are obtainable and stable as a result of the linker arrangement as shown in Figure 4 and schematically depicted in **IIIa** and **IVa** (Scheme 3).

CD spectra of **1–3** and  $[\text{d}(\text{TG}_4\text{T})]_4$  in  $\text{Na}^+$  buffer at 25 °C match those observed in  $\text{K}^+$  showing positive and negative bands at 262 and 244 nm, respectively, whereas in the CD profile of **4** in  $\text{Na}^+$  buffer a slight shift of the positive and negative bands at 259 and 239 nm, respectively, takes place. Furthermore, in the latter the presence of a new lower positive band centered at 289 nm is observable. These data indicate that even in  $\text{Na}^+$  buffer, possessing a lower stabilizing effect on G-quadruplex structures when compared to  $\text{K}^+$  buffer (3, 48), the TEL-ODN **1–3** are structured in parallel quadruplexes. For **4** in  $\text{Na}^+$  buffer we hypothesize the coexistence of two or more conformational species, among which the parallel (main) **IVa** and the antiparallel (minor) **IVb,c** structures are present, considering that most antiparallel G-quadruplexes are characterized by a positive band centered at 290–295 nm and a negative

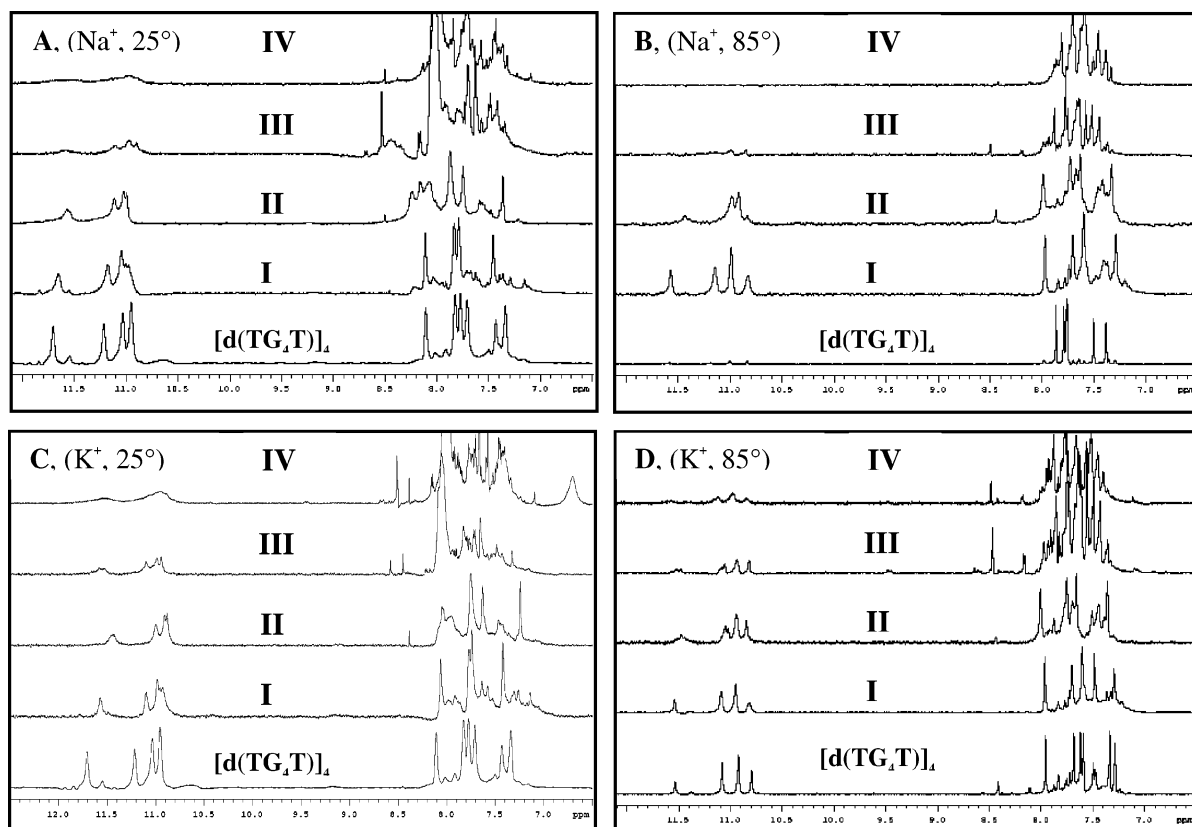


**Figure 2.** CD and CD melting spectra of **I–IV** and  $d(TG_4T)_4$  in  $K^+$  buffer (left column), in  $Na^+$  buffer (middle column) at 25 °C (solid line) and 90 °C (dotted line).  $CD_{265}$  melting profiles (right column) in  $K^+$  (solid line) and  $Na^+$  (dotted line).

band centered around 265 nm (44–47). CD spectra of **1–4** and  $d(TG_4T)_4$  at 90 °C for both buffers are reported in Figure 2. The data in  $K^+$  buffer confirm that the parallel G-quadruplex structures are completely retained in the complexes **I** and **II** while a detectable destructure can be observed for complexes **III** and **IV**. CD spectra in  $Na^+$  at 90 °C suggest that complexes **I–III** retain, to some extent the characteristic parallel G-quadruplex profile, while complex **IV** and  $d(TG_4T)_4$  are completely unstructured to random coils.

CD thermal denaturation–renaturation experiments were performed monitoring the CD value (mdeg) at 265 nm in the range 25–90 °C with 1 °C/min temperature scan rate. CD melting curves in  $K^+$  buffer for quadruplexes **I**, **II** and  $d(TG_4T)_4$  underwent irrelevant or low  $CD_{265}$  variations up to 90 °C while a relatively high  $CD_{265}$  variation was shown by the complexes **III** and **IV**. The melting profiles of the complexes **I–IV** and  $d(TG_4T)_4$  in  $Na^+$  showed higher  $CD_{265}$  variation with respect to the corresponding values registered in  $K^+$ ;





**Figure 3.**  $^1\text{H}$  NMR (500 MHz) spectra of **I–IV** and  $[\text{d}(\text{TG}_4\text{T})]_4$  in  $\text{Na}^+$  and  $\text{K}^+$  buffers. The spectra were registered in  $\text{H}_2\text{O}/\text{D}_2\text{O}$ , (9:1, v/v) solution at 0.5 mM of quadruplex at 25 and 85 °C. The aromatic and imino proton regions of the spectra are reported.

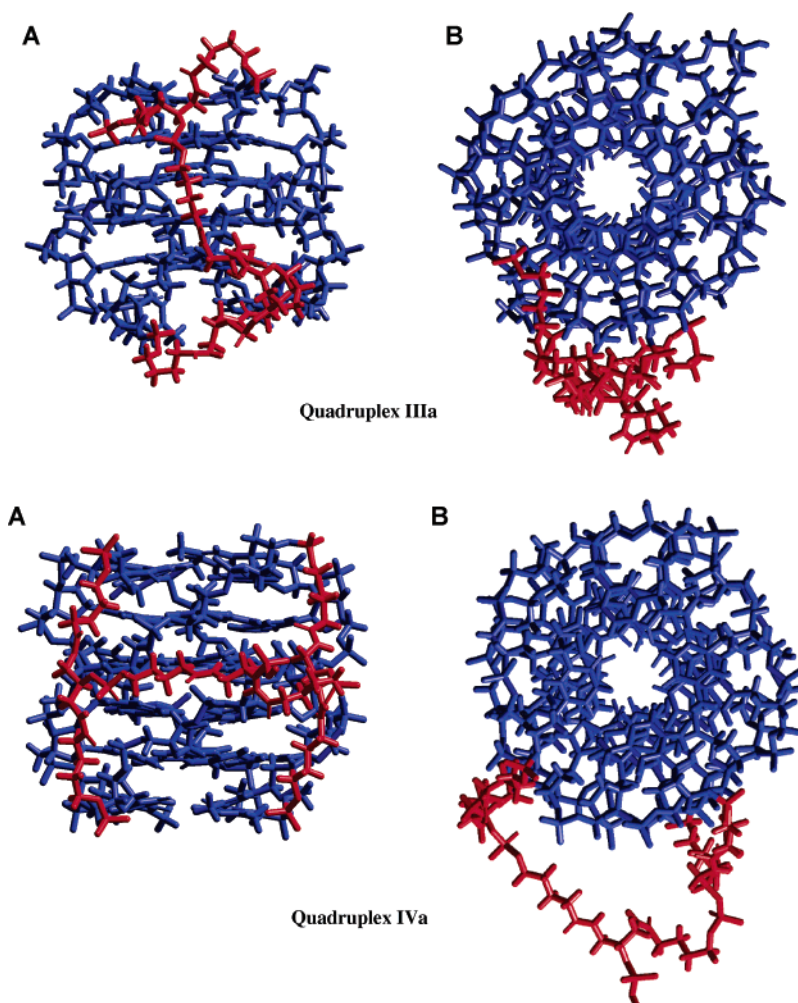
however, only for the  $[\text{d}(\text{TG}_4\text{T})]_4$  complex was a typical and derivatizable sigmoidal curve ( $T_m = 59\text{ }^\circ\text{C}$ ) registered. These data indicate that the melting process for the complexes **I–III** in  $\text{Na}^+$  is not complete at  $90\text{ }^\circ\text{C}$ , while a complete melting of **IV** can be observed in  $\text{Na}^+$  at this temperature. However, considering the shape of the apparent sigmoidal curve and the possible polymorphism of **IV**, we could not determine its melting temperature in both buffers. We also registered the  $\text{CD}_{265}$  thermal renaturation curves for all complexes in the range  $90\text{--}25\text{ }^\circ\text{C}$ . For the tetramolecular  $[\text{d}(\text{TG}_4\text{T})]_4$  complex, as expected, a very slow renaturation process and a high hysteresis phenomenon was indicated by the profile of the renaturation curve at  $1\text{ }^\circ\text{C}/\text{min}$  scan rate. The renaturation curves of **I** and **II** showed relatively low hysteresis at the same temperature scan rate. The fast reformation of the complexes **I** and **II** was also confirmed by the CD spectrum profiles registered at the end of the renaturation processes (data not shown). In the case of complexes **III** and **IV** the renaturation curves showed high hysteresis associated with low restructuring levels.

#### $^1\text{H}$ NMR Studies on Quadruplexes **I–IV** and $[\text{d}(\text{TG}_4\text{T})]_4$

$^1\text{H}$  NMR studies on quadruplexes **I–IV** were performed, after annealing, in the same buffers used for CD measurements. The spectrum of **I** and **II**, recorded in  $\text{Na}^+$  buffer at  $25\text{ }^\circ\text{C}$  (Figure 3A), showed the presence of four singlets in the range  $10.9\text{--}11.8$ , two of them partially overlapped, attributable to exchange protected imino protons involved in Hoogsteen  $\text{N}(1)/\text{O}(6)$  hydrogen bonds of G-quartets (45, 48, 49). These signals were almost superimposable with those observed in the spectrum of the  $[\text{d}(\text{TG}_4\text{T})]_4$  complex already characterized by NMR data (50). The  $^1\text{H}$  NMR spectra recorded in  $\text{Na}^+$  buffer at  $85\text{ }^\circ\text{C}$  (Figure 3B) indicated that **I** is completely stable at least up to  $85\text{ }^\circ\text{C}$ . **II** is partially preserved (only two imino proton signals are still detectable at this temperature), whereas the tetramolecular complex  $[\text{d}(\text{TG}_4\text{T})]_4$  is completely unstructured as indicated by the disappearance of all the imino proton signals.

Upon the basis of these data, it appears that quadruplex **II** in  $\text{Na}^+$  buffer is more stable than the tetramolecular counterpart, yet it is less stable than quadruplex **I**. The  $^1\text{H}$  NMR spectra of **I**, **II**, and  $[\text{d}(\text{TG}_4\text{T})]_4$  recorded in  $\text{K}^+$  buffer displayed quite similar signal patterns; however, it was not possible to observe any quenching of the imino proton signals up to  $85\text{ }^\circ\text{C}$  (Figure 3D). This observation is in accordance with reported data about the capability of  $\text{K}^+$  cations to form more stable quadruplex structures in comparison with those observed in  $\text{Na}^+$  buffer (51). As expected,  $^1\text{H}$  NMR spectra of **III** indicated that this compound is more structured in  $\text{K}^+$  than in  $\text{Na}^+$  buffer. In  $\text{K}^+$  buffer it was possible to detect the four imino proton signals up to  $85\text{ }^\circ\text{C}$  (Figure 3C and 3D), whereas in  $\text{Na}^+$  the spectrum showed large imino proton signals at  $25\text{ }^\circ\text{C}$  (Figure 3A) which almost disappeared at  $85\text{ }^\circ\text{C}$  (Figure 3B). The  $^1\text{H}$  NMR spectra of **IV** at  $25\text{ }^\circ\text{C}$  showed very large and overlapped imino proton signals both in  $\text{Na}^+$  (Figure 3A) and  $\text{K}^+$  buffers (Figure 3C). The increase of temperature for  $^1\text{H}$  NMR spectra of **IV** led to the sharpening of all imino proton signals both in  $\text{Na}^+$  and  $\text{K}^+$  buffers. In  $\text{K}^+$  buffer at least three imino proton signals are clearly detectable even at  $85\text{ }^\circ\text{C}$  (Figure 3D), whereas in  $\text{Na}^+$  buffer all imino proton signals are completely lost since  $65\text{ }^\circ\text{C}$  (data not shown). These data, which are consistent with UV and CD experiments, indicate that both complexes **III** and **IV** result in less stable quadruplex structures than **I** and **II**. Furthermore, **III** results more stable than **IV** especially in  $\text{K}^+$  buffer.

**Molecular Modeling.** To explore the capability of TEL-ODNs **3** and **4** of adopting the parallel quadruplex structure **IIIa** and **IVa**, respectively, a modeling study was performed. The minimized structures obtained are shown in Figure 4. Inspection of these structures reveals that in both **IIIa** and **IVa**, the tetra-end-linker is able to span the distance between the 5'- and 3'-ends of the strands assembled in parallel quadruplex structure. It can be noted that the orientation of the linker with respect to the helical axis is different in **IIIa** and **IVa**



**Figure 4.** Molecular model pictures of quadruplex structures **IIIa** and **IVa**; A: front view; B: top view. The tetra-end-linker is in red.

quadruplexes. For both the complexes the interatomic distances between the linker atoms and the quadruplex groove hydrogen bonding sites suggest the possibility of favorable nonbonded interactions. Although the integrity of the G-tetrads is maintained after energy minimization for both complexes, a detailed analysis of the H-bonds scheme inside the G-tetrads reveals some differences between the quadruplexes. In particular, in **IIIa**, each G-tetrad shows the same hydrogen bond scheme with an optimum distance and angle between donor and acceptor atoms, resulting in strong hydrogen bonds. Indeed in **IVa** there is a change of hydrogen-bond distances and angles between adjacent quartets, and, on average, two hydrogen bonds for each G-tetrad are lost. Furthermore, the degree of coplanarity between the G residues in each G-tetrad is greater in **IIIa** than in **IVa**. It is important to point out that the perturbation of each G-tetrad directly affects the stacking energy of the neighboring G-tetrads. In summary, the TEL-ODN **3** can form a parallel quadruplex structure **IIIa** without causing any distortion in the G-tetrad region, whereas in the TEL-ODN **4**, which forms the quadruplex **IVa**, a major readjustment of the G-tetrads occurs. These findings appear consistent with the experimental CD and NMR data revealing the formation of a more stable quadruplex **III** in comparison to **IV**.

The results described in this paper confirm the potential of TEL-ODNs as a new class of ODN analogues prone to adopt stable intramolecular quadruplex structures. In each case the  $T_m$  values are higher in potassium- than in sodium-containing buffer. It is noteworthy that, irrespective of the ends by which the four strands are attached to the tetra-end-linker, the CD spectra of the obtained complexes are typical of parallel

quadruplex structures, displaying positive maxima centered at 264 nm, in contrast to maxima at 295 nm observed for antiparallel structures. The only exception concerns the CD spectrum of **4** in  $\text{Na}^+$  buffer, in which a positive shoulder at 295 nm is present, indicating that a mixture of parallel and antiparallel forms could exist (quadruplex **IVa–c**). The whole of data, including molecular modeling calculations, suggest that the tetra-end-linker used in the present study is long and flexible enough to allow the organization of the ODN strands into a parallel quadruplex structure and (ii) that, at least in the presence of potassium ions, this form is favored over the antiparallel one, even if it is at the cost of a major readjustment of G-tetrads, which is the case of **IVa**. We believe that TEL-quadruplexes with suitable sequences and predetermined strand orientations could be used as aptamers, decoy and molecular probes, when used in a labeled form. Furthermore, TEL-ODNs may prove to be important tools for structural studies of alternative less stable quartets such as A-tetrads (52, 53), C-tetrads (54), T-tetrads (55), and mixed tetrads (56) as well as tetrads formed by modified nucleosides (57, 58). To confirm these conclusions as well as to get hold of a suitable tool to prepare monomolecular quadruplex structures with a preordered strand orientation, the syntheses of several TEL-ODNs characterized by shorter branched spacers and/or longer DNA stretches are currently in progress in our laboratory.

#### ACKNOWLEDGMENT

The authors are grateful to “Centro di Servizi Interdipartimentale di Analisi Strumentale”, C.S.I.A.S., for supplying NMR

facilities. This work is supported by Italian M.U.R.S.T. (P.R.I.N. 2003 and 2004) and Regione Campania (L.R. 5).

## LITERATURE CITED

- (1) Davis, J. T. (2004) G-quartets 40 years later: from 5'-GMP to molecular biology and supramolecular chemistry. *Angew. Chem., Int. Ed.* **43**, 668–698.
- (2) Keniry, M. A. (2001) Quadruplex structures in nucleic acids. *Biopolymers* **56**, 123–146.
- (3) Simonsson, T. (2001) G-Quadruplex DNA Structures Variations on a Theme. *Biol. Chem.* **382**, 621–628.
- (4) Arthanari, H., and Bolton, P. H. (2001) Functional and dysfunctional roles of quadruplex DNA in cells. *Chem. Biol.* **8**, 221–238.
- (5) Cuesta, J., Read, M. A., and Neidle, S. (2003) The Design of G-quadruplex ligands as telomerase inhibitors. *Mini Rev. Med. Chem.* **3**, 11–21.
- (6) Neidle, S., and Parkinson, G. (2002) Telomere maintenance as a target for anticancer drug discovery. *Nature Rev. Drug Discovery* **1**, 383–393.
- (7) Kerwin, S. M. (2000) G-Quadruplex DNA as a target for drug design. *Curr. Pharm. Des.* **6**, 441–466.
- (8) White, L. K., Wright, W. E., and Shay, J. W. (2001) Telomerase inhibitors. *Trends Biotechnol.* **19**, 114–120.
- (9) Ambrus, A., Chen, D., Dai, J., Jones, R. A., and Yang, D. (2005) Solution structure of the biologically relevant G-quadruplex element in the human c-MYC promoter. Implications for G-quadruplex stabilization. *Biochemistry* **44**, 2048–2058.
- (10) Seenisamy, J., Rezler, E. M., Powell, T. J., Tye, D., Gokhale, V., Joshi, C. S., Siddiqui-Jain, A., Hurley, L. H. (2004) The dynamic character of the G-quadruplex element in the c-MYC promoter and modification by TMPyP4. *J. Am. Chem. Soc.* **126**, 8702–8709.
- (11) Simonsson, T., Pecinka, P., and Kubista, M. (1998) DNA tetraplex formation in the control region of c-myc. *Nucleic Acids Res.* **26**, 1167–1172.
- (12) Macaya, R. F., Schultze, P., Smith, F. W., Roe, J. A., and Feigon, J. (1993) Thrombin-binding DNA aptamer forms a unimolecular quadruplex structure in solution. *Proc. Natl Acad. Sci. U.S.A.* **90**, 3745–3749.
- (13) Jing, N., Marchand, C., Guan, Y., Liu, J., Pallansch, L., Lackman-Smith, C., De Clercq, E., and Pommier, Y. (2001) Structure-activity of inhibition of HIV-1 integrase and virus replication by G-quartet oligonucleotides. *DNA Cell Biol.* **20**, 499–508.
- (14) Smirnov, I., and Shafer, R. H. (2000) Effect of loop sequence and size on DNA aptamer stability. *Biochemistry* **39**, 1462–1468.
- (15) Pileur, F., Andreola, M. L., Dausse, E., Michel, J., Moreau, S., Yamada, H., Gaidamarov, S. A., Crouch, R. J., Toulmé, J. J., and Cazenave, C. (2003) Selective inhibitory DNA aptamers of the human RNase H1. *Nucleic Acids Res.* **31**, 5776–5788.
- (16) Chinnappen, D. J., and Sen, D. (2002) Hemin-stimulated docking of cytochrome c to a hemin-DNA aptamer complex. *Biochemistry* **41**, 5202–5212.
- (17) Hammond-Kosack, M. C., Kilpatrick, M. W., and Docherty, K. (1992) Analysis of DNA structure in the human insulin gene-linked polymorphic region *in vivo*. *J. Mol. Endocrinol.* **9**, 221–225.
- (18) Delaney, S., and Barton, J. K. (2003) Charge Transport in DNA Duplex/Quadruplex Conjugates. *Biochemistry* **42**, 14159–14165.
- (19) de Champdore, M., De Napoli, L., Montesarchio, D., Piccialli, G., Caminal, C., Mulazzani, Q. G., Navacchia, M. L., and Chatgililoglu, C. (2004) Excess electron transfer in G-quadruplex. *Chem. Commun.* **15**, 1756–1757.
- (20) Calzolari, A., Di Felice, R., Molinari, E., and Garbesi, A. (2004) Electron Channels in Biomolecular Nanowires. *J. Phys. Chem. B* **108**, 2509–2515.
- (21) Parkinson, G. N., Lee, M. P. H., and Neidle, S. (2002) Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature* **417**, 876–880.
- (22) Patel, D. J. (2002) A molecular propeller. *Nature* **417**, 807–808.
- (23) Phan, A. T., Modi, Y. S., and Patel, D. J. (2004) Propeller-type parallel-stranded G-quadruplexes in the human c-myc promoter. *J. Am. Chem. Soc.* **126**, 8710–8716.
- (24) Ambrus, A., Chen, D., Dai, J., Jones, R. A., and Yang, D. (2005) Solution structure of the biologically relevant G-quadruplex element in the human c-myc promoter. Implications for G-quadruplex stabilization. *Biochemistry* **44**, 2048–2058.
- (25) Chou, S.-H., Chin, K.-H., and Wang, A. H.-J. (2005) DNA aptamers as potential anti-HIV agents. *Trends Biochem. Sci.* **30**, 231–234.
- (26) Koizumi, M., Akahori, K., Ohmine, T., Tsutsumi, S., Sone, J., Kosaka, K., Kaneko, M., Kimura, S., and Shimada, K. (2000) Biologically active oligodeoxyribonucleotides. Part 12: N2-methylation of 2'-deoxyguanosines enhances stability of parallel G-quadruplex and anti-HIV-1 activity. *Biorg. Med. Chem. Lett.* **10**, 2213–2216.
- (27) Wyatt, J. R., Vickers, T. A., Roberson, J. L., Buckheit, R. W., Jr., Klimkait, T., DeBaets, E., Davis, P. D., Rayner, B., Imbach, J. L., and Ecker, D. J. (1994) Combinatorially selected guanosine-quartet structure is a potent inhibitor of human immunodeficiency virus envelope-mediated cell fusion. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 1356–1360.
- (28) Wyatt, J. R., Davis, P. W., and Freier, S. M. (1996) Kinetics of G-quartet-mediated tetramer formation. *Biochemistry* **35**, 8002–8008.
- (29) Hazel, P., Huppert, J., Balasubramanian, S., and Neidle, S. (2004) Loop-length-dependent folding of G-quadruplexes. *J. Am. Chem. Soc.* **126**, 16405–16415.
- (30) Risitano, A., and Fox, K. R. (2004) Influence of loop size on the stability of intramolecular DNA quadruplexes. *Nucleic Acid Res.* **32**, 2598–2606.
- (31) Jing, N., Rando, R. F., Pommier, Y., and Hogan, M. E. (1997) Ion selective folding of loop domains in a potent anti-HIV oligonucleotides. *Biochemistry* **36**, 12498–12505.
- (32) Rujian, I. L., Meleney, J. C., and Bolton, P. H. (2005) Vertebrate telomere repeat DNAs favour external loop propeller quadruplex structures in the presence of high concentrations of potassium. *Nucleic Acid Res.* **33**, 2022–2031.
- (33) Risitano, A., and Fox, K. R. (2003) Stability of intramolecular DNA quadruplexes: comparison with DNA duplexes. *Biochemistry* **42**, 6507–6513.
- (34) Lu, M., Guo, Q., and Kallenbach, N. R. (1993) Thermodynamics of G-tetraplex formation by telomeric DNAs. *Biochemistry* **32**, 598–601.
- (35) Zhou, T., Chen, G., Wang, Y., Zhang, Q., Yang, M., and Li, T. (2004) Synthesis of unimolecular circular G-quadruplexes as prospective molecular probes. *Nucleic Acid Res.* **32**, e173/1–e173/9.
- (36) Oliviero, G., Borbone, N., Galeone, A., Varra, M., Piccialli, G., and Mayol, L. (2004) Synthesis and characterization of a bunched oligonucleotide forming a monomolecular parallel quadruplex structure in solution. *Tetrahedron Lett.* **45**, 4869–4872.
- (37) Oliviero, G., Amato, J., Borbone, N., Galeone, A., Varra, M., Piccialli, G., Mayol, L. (2006) Synthesis and characterization of DNA quadruplexes containing T-tetrads formed by bunch-oligonucleotides. *Biopolymers* **81**, 194–201.
- (38) Breslauer, K. J., Frank, R., Blöcker, H., and Marky, L. A. (1986) Predicting DNA duplex stability from the base sequence. *Proc. Natl. Acad. Sci. U. S. A.* **83**, 3746–3750.
- (39) Piotto, M., Saudek, V., and Sklenar, V. (1992) Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J. Biomol. NMR* **2**, 661–665.
- (40) Cornell, W. D., Cieplack, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. (1995) A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **117**, 5179–5197.
- (41) Gasteiger, J., and Marsili, M. (1980) Iterative partial equalization of orbital electronegativity: a rapid access to atomic charges. *Tetrahedron* **36**, 3219–3228.
- (42) Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **79**, 926–935.
- (43) Nielsen, J., and Lyngsø, L. O. (1996) Combinatorial Solid-Phase Synthesis of Balanol Analogues. *Tetrahedron Lett.* **37**, 8439–8442.
- (44) Petraccone, L., Erra, E., Esposito, V., Randazzo, A., Mayol, L., Nasti, L., Barone, G., and Giancola, C. (2004) Stability and structure of telomeric DNA sequences forming quadruplexes containing four G-tetrads with different topological arrangements. *Biochemistry* **43**, 4877–4884.
- (45) Jin, R., Gaffney, B. L., Wang, C., Jones, R. A., and Breslauer, K. J. (1992) Thermodynamics and Structure of a DNA Tetraplex: A

- Spectroscopic and Calorimetric Study of the Tetramolecular Complexes of d(TG<sub>3</sub>T) and d(TG<sub>3</sub>T<sub>2</sub>G<sub>3</sub>T). *Proc. Natl. Acad. Sci. U. S. A.* 89, 8832–8836.
- (46) Dapić, V., Abdomerović, V., Marrington, R., Pederby, J., Rodger, A., Trent, J. O., and Bates, P. J. (2003) Biophysical and biological properties of quadruplex oligodeoxyribonucleotides. *Nucleic Acids Res.* 31, 2097–2170.
- (47) Hardin, C. C., Perry, A. G., and White, K. (2001) Thermodynamic and kinetic characterization of the dissociation and assembly of quadruplex nucleic acids. *Biopolymers* 56, 147–194.
- (48) Feigon, J., Koshlap, K. M., and Smith, F. W. (1995) <sup>1</sup>H NMR spectroscopy of DNA triplexes and quadruplexes. *Methods Enzymol.* 261, 225–255.
- (49) Feigon, J. (1996) DNA Triplexes, Quadruplexes, and Aptamers. *Encycl. Nuclear Magn. Reson.* 3, 1726–1731.
- (50) Aboul-ela, F., Murchie, A. I. H., Norman, D. G., and Lilley, D. M. J. (1994) Solution structure of a parallel-stranded tetraplex formed by d(TG<sub>4</sub>T) in the presence of sodium ions by nuclear magnetic resonance spectroscopy. *J. Mol. Biol.* 243, 458–471.
- (51) Hud, N. V., Smith, F. W., Anet, F. A., and Feigon, J. (1996) The selectivity for K<sup>+</sup> versus Na<sup>+</sup> in DNA quadruplexes is dominated by relative free energies of hydration: a thermodynamic analysis by <sup>1</sup>H NMR. *Biochemistry* 35, 15383–15390.
- (52) Patel, P. K., Koti, A. S. R., and Hosur, R. V. (1999) NMR studies on truncated sequences of human telomeric DNA: observation of a novel A-tetrad. *Nucleic Acids Res.* 27, 3836–3843.
- (53) Gavathiotis, E., and Searle, M. S. (2003) Structure of the parallel-stranded DNA quadruplex d(TTAGGGT)<sub>4</sub> containing the human telomeric repeat: evidence for A-tetrad formation from NMR and molecular dynamics simulations. *Org. Biomol. Chem.* 1, 1650–1656.
- (54) Patel, P. K., Bhavesh, N. S., and Hosur, R. V. (2000) NMR observation of a novel C-tetrad in the structure of the SV40 repeat sequence GGGCGG. *Biochem. Biophys. Res. Commun.* 270, 967–971.
- (55) Patel, P. K., and Hosur, R. V. (1999) NMR observation of T-tetrads in a parallel stranded DNA quadruplex formed by *Saccharomyces cerevisiae* telomere repeats. *Nucleic Acids Res.* 27, 2457–2464.
- (56) Meyer, M., Schneider, C., Brandl, M., and Suehnel, J. (2001) Cyclic Adenine-, Cytosine-, Thymine- and Mixed Guanine-Cytosine-Base-Tetrads in Nucleic Acids Viewed From a Quantum-Chemical and Force Field Perspective. *J. Phys. Chem. A* 105, 11560–11573.
- (57) Esposito, V., Randazzo, A., Piccialli, G., Petraccone, L., Giancola, C., and Mayol, L. (2004) Effects of an 8-bromodeoxyguanosine incorporation on the parallel quadruplex structure [d(TGGGT)]<sub>4</sub>. *Org. Biomol. Chem.* 2, 313–318.
- (58) Chen, J., Zhang, L. R., Min, J. M., and Zhang, L. H. (2002) Studies on the synthesis of a G-rich octaoligonucleotide (isoT)<sub>2</sub>(isoG)<sub>4</sub>-(isoT)<sub>2</sub> by the phosphotriester approach and its formation of G-quartet structure. *Nucleic Acids Res.* 30, 3005–3014.

BC060009B



# Synthesis and Characterization of Monomolecular DNA G-Quadruplexes Formed by Tetra-End-Linked oligonucleotides

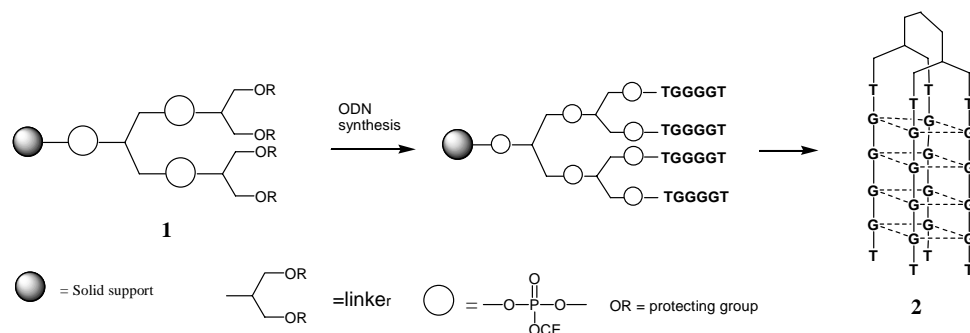
**Jussara Amato**

*Dipartimento di Chimica delle Sostanze Naturali,  
Università di Napoli Federico II Facoltà di Scienze Biotecnologiche  
via D. Montesano 49, I-80131 Napoli Italy  
jussara.amato@unina.it*

**ABSTRACT:** Guanine-rich DNA sequences are widely dispersed in eukaryotic genome and are abundant in regions with relevant biological significance. They can form quadruplex structures from stacked tetrads of hydrogen-bonded guanine bases. These stable, non classical, secondary structures can be intermolecular or intramolecular in nature and can exhibit structural polymorphism based on variability in both strand polarity and loop geometry. However, intermolecular formation of quadruplex structure in vitro is very slow and may require high ODN concentrations. These unfavourable kinetic and thermodynamic parameters could be disadvantageous in view of their potential therapeutic use. We report here the syntheses and the structural studies of a set of interconnected d(TG<sub>4</sub>T) fragments which are tethered, with different orientations, to a tetra-end-linker in an attempt to force the formation of specific four-stranded DNA quadruplex structures. Two synthetic strategies have been used to obtain oligodeoxyribonucleotide (ODN) strands linked with their 3' or 5' ends to each of the four arms of the linker. The first approach allowed the synthesis of tetra end-linked ODN (TEL-ODN) containing the four ODN strands with a parallel orientation, while, the latter synthetic pathway led to the synthesis of TEL-ODNs each containing antiparallel ODN pairs. The influence of the linker at 3' or 5'-ODN, on the quadruplex typology and stability, in the presence of sodium or potassium ions, has been investigated by circular dichroism (CD), CD thermal denaturation, <sup>1</sup>H NMR experiments at variable temperature and molecular modeling. All synthesized TEL-ODNs showed to form parallel G-quadruplex structures. Particularly, the TEL-ODN containing all parallel ODN tracts formed very stable parallel G-quadruplex complexes whereas, the TEL-ODNs containing antiparallel ODN pairs led to relatively less stable parallel G-quadruplexes. The molecular modeling data suggested that the above antiparallel TEL-ODNs can adopt parallel G-quadruplex structures thanks to a considerable folding of the tetra-branched-linker around the whole quadruplex scaffold. The synthesis of a new linker with a shorter branched-spacers and or longer DNA stretches are currently in progress in our laboratory. We believe that TEL-quadruplexes with suitable



sequences and predetermined strand orientations could be used as aptamers, decoy and molecular probes when used in a labelled form. Furthermore, TEL-ODNs may prove to be important tools for structural studies of alternative less stable quartets such as A-tetrads, C-tetrads, T-tetrads mixed tetrads as well as tetrads formed by modified oligonucleotides.



### Riferimenti bibliografici:

- Lane, A. N.; Jenkins, T. C. *Curr. Org. Chem.* **2001**, 5, 845-869.
- Belmont, P.; Constant, J. F.; Demeunynck, M. *Chem. Soc. Rev.* **2001**, 30, 70-81.
- Shafer, R. H.; Smirnov, I. *Biopolymers* **2001**, 56, 209-227.
- Jing, N.; Marchand, C.; Guan, Y.; Liu, J.; Pallansch, L.; Lackman-Smith, C.; De Clercq, E.; Pommier, Y. *DNA Cell Biol.* **2001**, 20, 499-508.
- Pileur, F.; Andreola, M.-L.; Dausse, E.; Michel, J.; Moreau, S.; Yamada, H.; Gaidamarov, S. A.; Crouch, R. J.; Toulmé, J. J.; Cazenave, C. *Nucleic Acids Res.* **2003**, 31, 5776-5788.
- Risitano, A.; Fox, K.R. *Nucleic Acids Res.* **2004**, 32, 2598-2606.
- Oliviero, G.; Borbone, N.; Galeone, A.; Varra, M.; Piccialli, G.; Mayol, L., *Tetrahedron Lett* **2004**, 45, 4869-4872.
- Rujian, I.L., Meleney, J.C., *Nucleic Acids Res.* **2005**, 33, 2022-2031